To Be but Not to Ski: Cognitive Modelling of Decision Making in a Ski Touring Context

Exam in Data Science, Prediction, and Forecasting

Kristian Severin Mengel-Niemann (201904957)

MSc Cognitive Science, School of Communication and Culture
University of Aarhus, Jens. Chr. Skous Vej 2, 8000 Aarhus C, Denmark
Lecturer: Chris Mathys
June 1st, 2023

AARHUS UNIVERSITY

All data and analyses-code can be found on:

https://github.com/kristianseverin/DataScienceSkiDecisionMaking.git

Introduction

1.1 Risk Aversion and Prospect Theory

Human risk-propensity is believed to be inherently risk-aversive (Dan & Loewenstein, 2020; O'Donoghue & Somerville, 2018). This belief originates from prospect theory where risk-propensity can be thought of in terms of how the prospect of uncertainty in a decision is believed to materialize itself for the individual making a choice, i.e., whether the decision yields possible gains or losses (Tversky & Kahneman, 1979, 1989). Prospect theory is traditionally applied to econometric contexts but has been shown to be widely applicable to a whole range of other domains such as choice-behavior in travelling (van de Kaa, 2010), expected utility represented as expected years of living as a function of health-choices (Attema et al., 2013), and in explaining political science models (McDermott, 2004).

1.2 Evolutionary Advantages of Risk-Aversion and Individual Variability

The study at hand builds on prospect-theory ideas where negative losses are felt more intensely than positive gains (Weller et al., 2011). More specifically, the individual variability that exists in loss-aversive behavior is of interest. In a paper on the evolutionary mechanisms introduced by stochasticity across generations Zhang et al. (2014) claim that risk-aversive behavior has inherent evolutionary benefits across generations and species. Furthermore, they use a simple binary-choice model to highlight individual differences in choice-behavior and link the variability to stochastic patterns of natural selection. This paper is particularly concerned with the individual variance in choices made in a backcountry skiing context wherein a binary choice yields the highest possible loss, i.e., death, and merely poses recreational fun as the potential gain.

Multiple studies have demonstrated that humans differ in their decision-making outcomes when deciding whether to ski tour. One study used an adaptation of Bayes' theorem to calculate probabilities of triggering an avalanche (Ebert, 2019). Ebert highlights that people vary in their decisions regardless of the avalanche-risk presented to them. Furthermore, Ebert argues that irrespective of the fact that Bayesian reasoning is a great tool for predicting the probability of triggering an avalanche, the introduction of it in avalanche-education curricula should be done with extreme caution. This is because Bayes' theorem is unintuitive in its nature, and people have demonstrated an inability in using it correctly (Westbury, 2010). A study on people's integration of heuristic information concluded that people are often predetermined to ski prior to doing so regardless of contradictory information from the ski-environment (Michaelsen & Rolland, 2016). Furthermore, habits were heavily predictive of where people were going to ski. People are inclined to have certain routes that they decide on skiing prior to entering the mountain. Thus, neglecting to incorporate dynamic information from the environment they are skiing in. Other papers have shown the importance, however, of utilizing multiple predictive elements to achieve safe skiing. A two-part paper by McClung (2002a, 2002b) attempted to incorporate both physical and human issues in avalanche forecasting. He argues that avalanche prevention presents itself as a dynamic problem constituted of both geotechnical predictors and human elements. He goes on to say that most avalanches are caused by human triggers. He divides avalanche prediction into seven subcomponents, among others, decision making. The paper at hand attempts to build on these findings and is primarily concerned with the human variability on self-informed decisions in a skiing context.

1.3 Individuals who Partake in Extreme Sports

Skiing is an extreme sport, and especially backcountry skiing in uncontrolled terrain. A meta-study on avalanche fatalities showed that around 100 people lose their lives in avalanche accidents in the European Alps each year (Techel et al., 2016). These fatalities should be viewed in the context of the extreme sport setting they occur in. A study using virtual reality linked personality to accident handling and found patterns between exhibiting hazardous behavior and personality (Ju et al., 2019). The researchers found that subjects who engaged in dangerous driving behavior scored higher on a range of personality-traits related to impulsiveness and goal-orientated behavior. The findings from this study adds to the hypothesis that certain types of personalities will have pre-destined motivations for their decision making in dangerous settings.

A logical assumption is consequently that people engaging in extreme sports, hereunder backcountry skiing, would share some personality traits that would make them likely to choose similarly. However, one study found that BASE-jumpers share only one trait, namely their low levels of harm-avoidance (Monasterio et al., 2012). In all other aspects, they are as different in personalities as a comparable control group. This is grounds for arguing that differences in personalities, even in groups of people who actively seek out dangerous activities, could be highly explanatory of decision making in dangerous contexts. In fact, if differences in personalities exist in a group of people who undertake even riskier activities than ski-touring, i.e., BASE-jumping, to an extent that is comparable with a control group, it would be logical to think that the activity itself does not attract only one specific type of personality. The author of the current paper therefore argues that conducting a study on differences in risk-propensity and how they influence decision-making in a ski-touring context is highly relevant. This is because knowledge about which personality-types are most likely to put themselves in knowingly dangerous settings, could provide mountaineering educators with personality-tailored educational tools.

Over the course of a ski season two skiers were tasked with risk-assessment of skiable terrain and to report whether they opted to ski the terrain. The risk-assessment followed a pre-made avalanche evaluation system. Cognitive modelling was done to capture proposed cognitive parameters in risk-assessment and risk-propensity. The nature of the study was purely proof-of-concept. The study was conducted with the intention of showing that individual differences underlying decision making can be explored using Bayesian cognitive modelling.

Methods

2.1 Participants

Two individuals participated in the proof-of-concept study. Both were males and were 23 and 24 years of age respectively. Both participants were expert skiers and ski-touring experts. Written informed consent was obtained before participation to comply with the Declaration of Helsinki. The participants reported 13 and 25 skiing assessments in 3 different countries between the 28th of February 2023 and the 6th of May 2023.

2.2 Experimental Design

The study was conducted using The Avaluator™ Version 2 slope evaluation card developed by Avalanche Canada (*The Avaluator*, 23/05/2023). Participants were asked to assess various avalanche conditions and terrain characteristics to get a two-dimensional score that would place their decision to ski/not to ski within a two-dimensional risk-evaluation space. Both the avalanche conditions and the terrain characteristics were comprised of sub-categories all of which the participants were asked to report as binary scores (0 if the phenomenon was non-present, 1 if it was present). The sub-categories and their descriptions can be seen in figure 1 similarly to how they were presented to the participants.

Avalanche Conditions	Score if phenomenon is present	Terrain Characteristics	Score if phenomenon is present
Regional Danger Rating: Is the avalanche danger rating "Considerable" or higher	+1	**Slope Steepness**: Is the slope steeper than 30 degrees?	+1
Persistent Avalanche Problem: Is there a persistent or deep persistent slab problem in the snowpack	+1	**Terrain Traps**: Are there gullies, trees, or cliffs that increase the consequences of being caught in an avalanche?	+1
Slab Avalanches: Are there signs of slab avalanches in the area from today or yesterday?	+1	**Slope shape**: is the slope convex or unsupported.	+1
Signs of instability: Are there signs of snowpack instability including *whumps*, shooting cracks, or drum-like sounds?	+1	Forest Density: Is the slope in the alpine, in a sparsely treed area or in open forest (cut-block, burn, wide-spaced glades)?	+1
Recent Loading: Has there been loading within the past 48 hours including roughly 30 cm of new snow or more, significant wind transport or rain?	+1		
Critical Warming: Has there been a recent rapid rise in temperature to near 0 C, or is the upper snowpack wet due to strong sun, above-freezing air temperatures or rain?	+1		
Avalanche Conditions Score	Accumulation of all avalanche conditions scores.	**Terrain Characteristics Score**	Accumulation of all terrain characteristics scores.

Figure 1, shows the Slope Evaluation card. Avalanche conditions score for a given run ranges from 0-6, and terrain charcteristics score ranges from 0-4. Note that for the version of the Avaluator™ provided by Avalanche Canada the characteristic "Slope steepness"

As seen from figure 1, the avalanche conditions score for a given run was an integer between 0 and 6, and the terrain characteristics score was an integer between 0 and 4. The Avaluator™ System gives the skier a dynamical hint of the riskiness associated with a decision. This is done by placing the overall evaluation of a given run within a two-dimensional space where a decision to ski a run can be accompanied by “Caution”, “Extra Caution”, or “Not Recommended”. The Avalanche conditions score is illustrated on the y-axis of the Avaluator™. Higher values see the avalanche conditions assessment be placed in *low, moderate, considerable, high, or extreme* were *low* covers avalanche conditions scores of 0 and 1 and each additional increment means moving to a higher risk assessment. The terrain complexity score runs along the x-axis and moves incrementally from lower risk to higher risk. Thus, the terrain can be assessed as *simple, challenging, or complex* with *simple* covering terrain characteristics scores of 0 and 1. Visit Avalanche Canada’s webpage to see an illustration of the copyrighted Avaluator™ system (avalanche.ca).

The slope evaluation does not give the skier a definitive answer as to whether a run is safe to ski or not. It serves as an assistive tool in the decision-making process. The participants were asked to report whether they chose to ski or whether they chose not to ski along with their avalanche conditions and terrain characteristic assessments.

2.3 Justification of Methodological Approach

In its core, the aim of this paper was to infer strategies or underlying mechanisms from decision-making behavior in a skiing context. The obstacles tackled by the researcher to illuminate these mechanisms consequently presented themselves as specific cases of ‘the inverse problem’ in which what the researcher has at their disposal is overt observations, and what the researcher is interested in is the underlying hidden causes that drive the behavior responsible for those overt observations. Two cases of Bayesian Cognitive models were implemented on the avalanche decision making data to estimate parameters inferred from the observed choices. Both models were hierarchical and were modelled using “brms” an R package for Bayesian Multilevel models using Stan (Bürkner, 2017; Stan Development Team. 2023).

1.4 Simple Bayes Model

A hierarchical bayes model was implemented to allow for multiple sources of information to impact the outcome of skiing or not skiing *O* on a given run *r* for a given skier *s*. The bayes model was denoted as:

$$logit\left(P\left(O_{r,s}\middle|a_{r,s},t_{r,s)}\right)\right) = logit(\beta_s) + logit\left(O_{r,s}\middle|a_{r,s}\right) + logit\left(O_{r,s}\middle|t_{r,s)}\right)$$

In which β is a parameter called *bias* to be estimated by the model, *a* is a source of information obtained by the skier, namely the avalanche conditions observed on a given run, and *t* is another source of information obtained by the skier, namely the terrain conditions for a given run. The model thus attempts to estimate a β-parameter from two different sources of information and the given outcomes for the combination of different values of those sources of information. The model is henceforth referred to as the "simple bayes model".

2.5 Weighted Bayes Model

An adaptation of the hierarchical simple bayes model with weights for the sources of information was implemented (Jardri et al., 2017). This model followed the same logic as the simple bayes model, with the important difference that it allowed for the two sources of information to be weighted relative to one another, i.e., two additional parameters were introduced for the model to estimate. A weight for the first source of information, *avalanche conditions (a),* was implemented as ω_1, while ω_2 was implemented as a weight parameter for the second source of information *terrain characteristics (t)*. The whole model was given by:

$$logit\left(P\left(O_{r,s}\middle|a_{r,s}, t_{r,s)}\right)\right) = logit(\beta_s) + logit\left(O_{r,s}\middle|a_{r,s}\right) * \omega_1 + logit\left(O_{r,s}\middle|t_{r,s)}\right) * \omega_2$$

2.6 Simulating data

It is common practice to validate a statistical, Bayesian model by simulating data from the proposed experimental setup before collecting data. By doing so, the researcher is allowed to test the model on data that is completely known to them and thereby check if the model produces results that should be expected from the scientific reasoning gone into the formulation of the model. 10 agents were simulated to ski with different bias-, weight1-, and weight2- levels for different skiing outcomes and values for the two sources of information. The data was passed to the model making it possible to see if the parameters could be successfully recovered for both realistic values, but also for more extreme cases. The data was simulated to most accurately mimic the decisions made by the skiers in the empirical data. Detailed information about the data simulation can be found in the GitHub repository linked at the start of this paper.

2.7 Model Quality Checks

Both models were validated with iterative quality checks. See appendix A for quality checks for the simple bayes model and Appendix B for quality checks for the weighted bayes model. Prior posterior update checks, prior posterior predictive checks, Hamiltonian Monte Carlo chain mixing checks, and prior sensitivity analysis were all done iteratively for both models when choosing priors for the parameter models, and when assessing what influence different prior values would have on the model estimations.

2.8 Parameter Recovery Simple Bayes

To test if the model produced results that captured bias values for both realistic and rare cases, parameter recovery was done. In effect, the proposed parameter values from the simulated skiers were compared to the estimated values from the model outputs. Bias was recovered successfully, confirming the mathematical reasoning that had gone into simulating the data and into the models themselves. Figure 2 shows the parameter recovery for bias from the simple bayes model. The figure illustrates that the simulated values and estimated values follow a linear relationship. Although, the relationship is not perfect, the relationship is deemed prevailing enough that bias was successfully recovered.

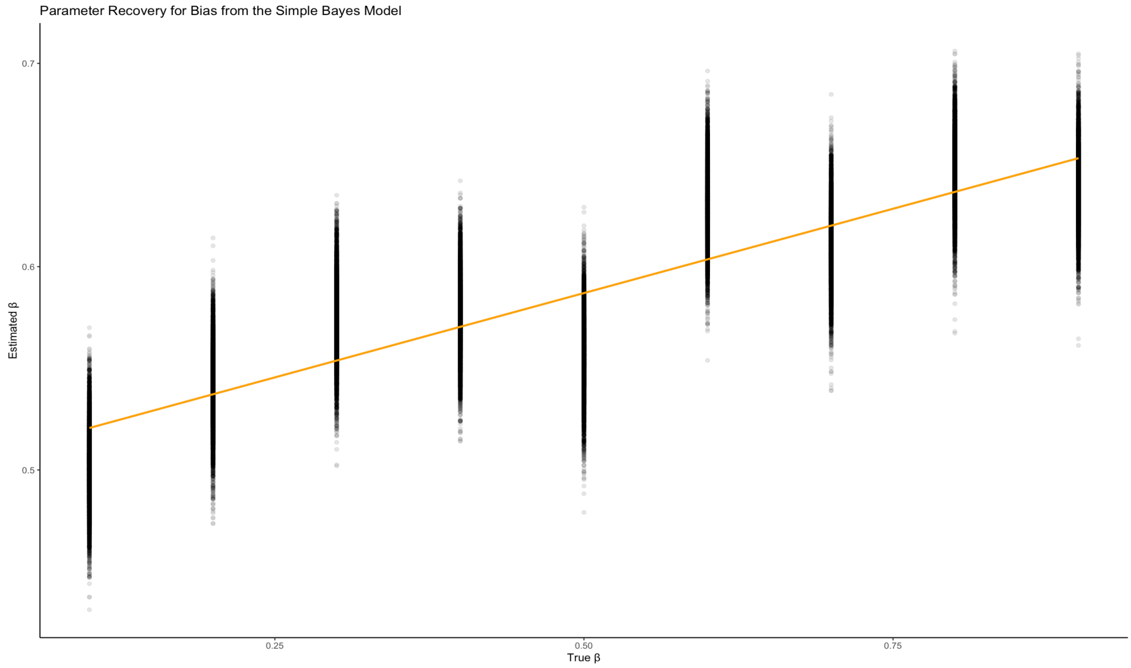

Figure 2, shows the hypothesized true bias on the x-axis and the estimated bias on the y-axis. The orange line is a linear regression fitted line.

2.9 Parameter Recovery Weighted Bayes

Similarly, parameter recovery was done on all the parameters from the weighted bayes model. All parameters were recovered successfully. Figure 3 shows the parameter recovery for the parameters in the weighted bayes model.

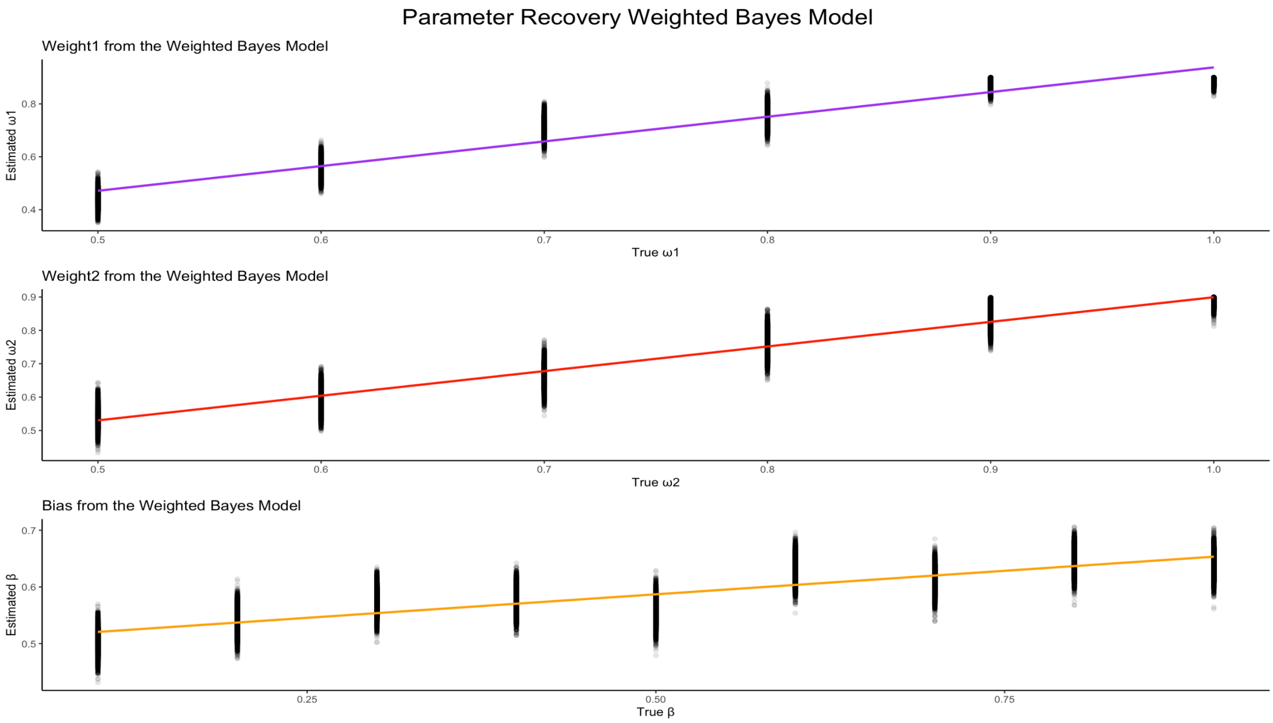

Figure 3, shows the estimated and hypothesized parameter values for all parameters in the weighted bayes model. The estimated values are shown on the y-axes, and the hypothesized, simulated values are shown on the x-axes. The colored lines are linear regression fitted lines.

Results

3.1 Simple Bayes Model

Both participants displayed a bias for skiing with 95% credible intervals above 0.5 (see table 2 for summary output). There was, however, a noticeable difference in their posterior bias distributions. Subject A had a mean that was considerably higher than the hypothesized naive true mean, i.e., 0.5 versus 0.75 (95% CI[0.59, 0.85]) while subject B had a mean that was even higher, i.e., 0.83 (95% CI[0.66, 0.95]). This suggests that there is a difference in the decision-making process of choosing to ski or not irrespective of consciously thought-out assessments of the risk involved in choosing to ski. Subject A is more risk-aversive than subject B that has a higher bias towards choosing to ski. Prior posterior update plots for bias can be seen in figure 4.

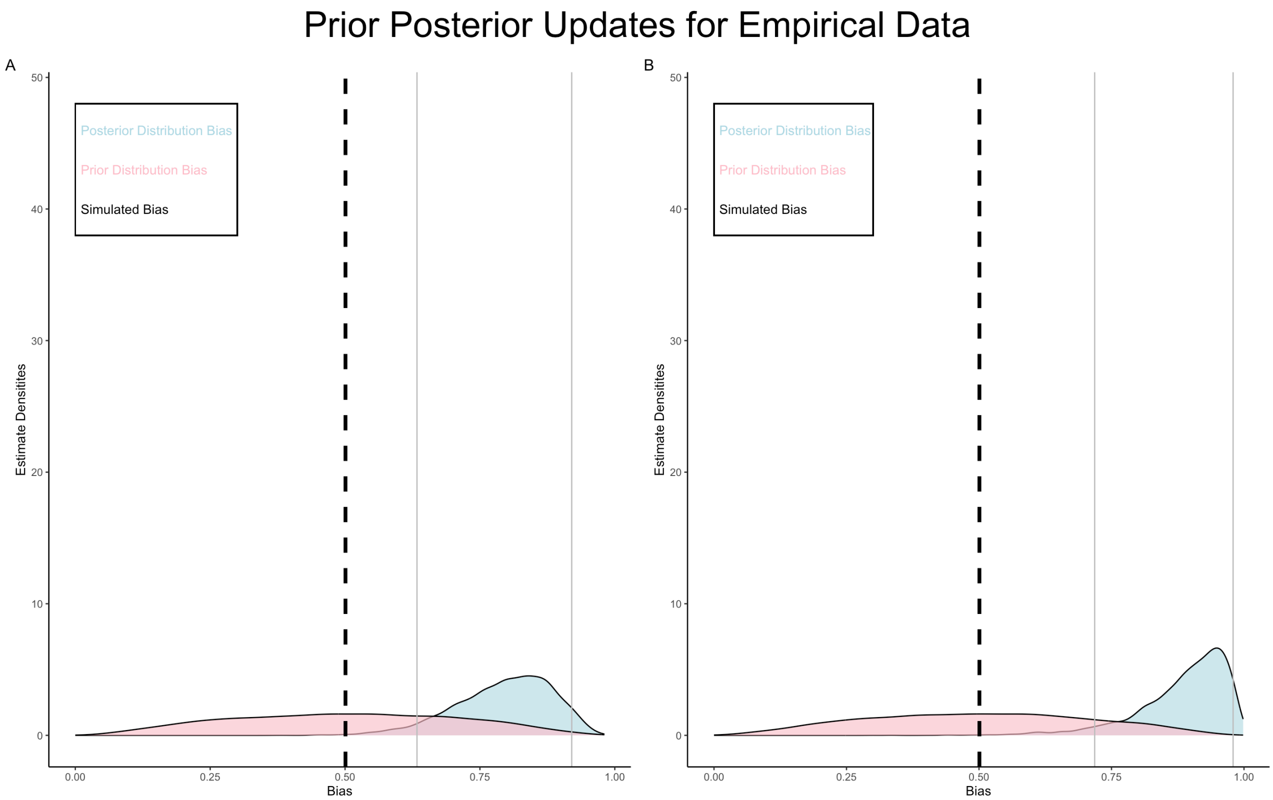

Figure 4, shows density distributions for the priors and posteriors for the bias parameter in the simple bayes model. The pink distributions are normally distributed around 0.5 and display the prior distributions. The blue distributions illustrate the posterior distributions for the estimated posteriors for the bias parameter for the two subjects. The dashed, black line shows the naive hypothesized bias of 0.5. The grey lines show the 95% credible intervals. The A and B refer to the subjects in the study.

	Mean	Median	Standard Deviation	MAD	q5	q95	rhat	ess_bulk	ess_tail
Bias Subject A	0.73	0.74	0.08	0.08	0.59	0.85	1.0001	6355	4325
Bias Subject B	0.83	0.84	0.09	0.08	0.66	0.95	1.0004	6102	4603

Table 1, shows summary output for the two subjects for the bias-parameter.

3.2 Weighted Bayes Model

The differences in the bias values between the two subjects became considerably less obvious for the weighted bayes model compared to the simple bayes model. The mean β-value for subject A was 0.7(95% CI[0.57, 0.82]), while it was 0.73 (95% CI[0.55, 0.88]) for subject B. Both subjects, however, displayed a far greater bias towards skiing than the hypothesized naive mean of 0.5. Subject A had a mean ω_1-value of 0.59 (95% CI[0.54, 0.66]), while it was 0.6 (95% CI[0.54, 0.67]) for subject B. The ω_2-parameter was also quite alike for the two subjects with mean-values of 0.58(95% CI[0.53, 0.64]) and 0.56 (95% CI[0.53, 0.62]) respectively.

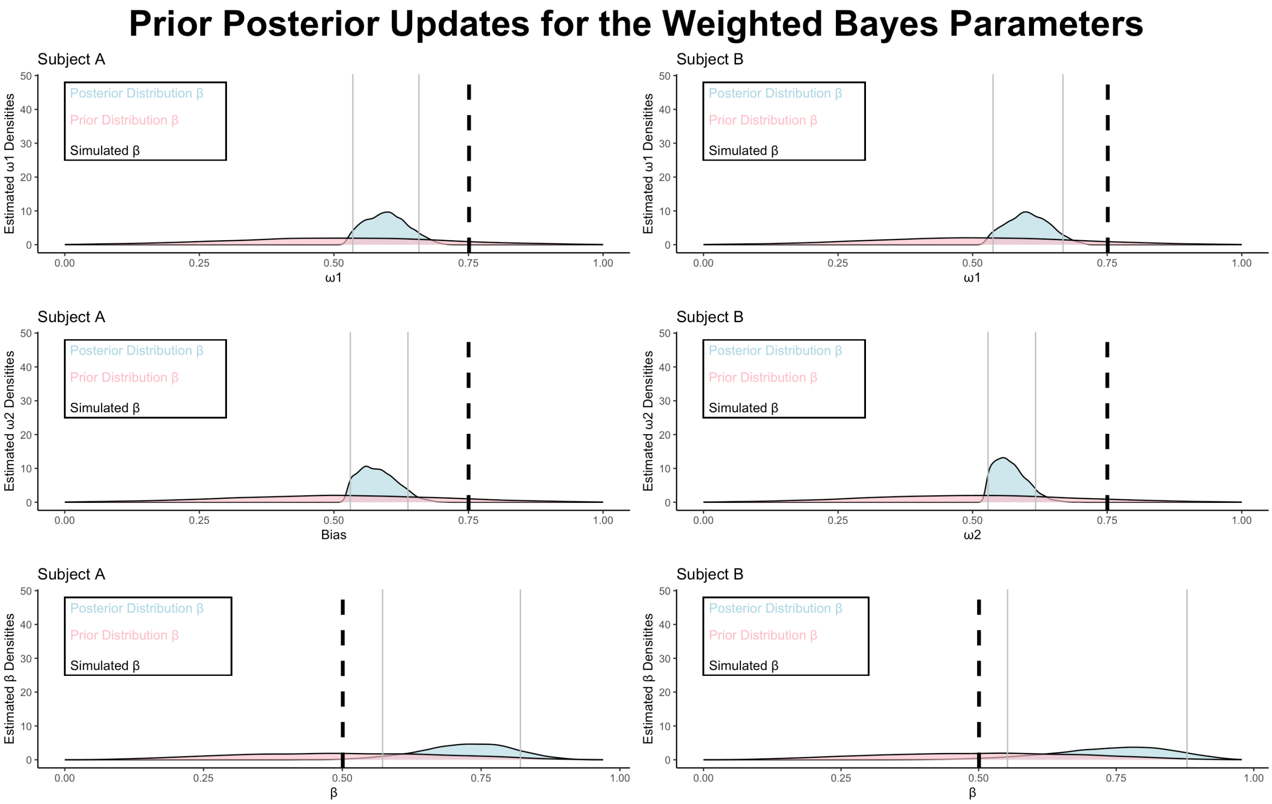

Figure 5, shows densities for the priors and posteriors for the bias, weight1, and weight2 parameters in the simple bayes model. The pink distributions are normally distributed around 0.5 and display the prior distributions. The blue distributions illustrate the posterior distributions for the estimated posteriors for the bias parameter for the two subjects. The dashed lines show the naive hypothesized parameter values. The grey lines show the 95% credible intervals for the parameter values.

Looking at figure 5, it looks like the two subjects had similar parameter distributions for all the parameters. However, the pattern of subject B being more biased towards skiing that was seen in the first model is still somewhat visible in the weighted bayes model. The tails of the posterior distribution seem somewhat longer for subject B, indicating more fluctuation for this subject’s bias between decisions. The posterior distributions also become much wider in the weighted bayes model. This could both be interpreted as the model not learning the true distributions for the subjects, or as the subjects displaying a natural wide range of bias from ski-run to ski-run. The distributions for both weight parameters seem to be estimated quite well and quite similarly for the two participants. This is also to be somewhat expected as the participants were acting based on the same danger-assessment system that would provide the subjects with guidance on how to explicitly weigh the two sources of information. Looking at the weight 2 parameter, it seems like a slight pareto-like pattern governs this weight for both participants. This could indicate some sort of threshold for which the accumulated terrain characteristics score would see the skiers decide to ski and everything above this score would have the participants ski with less and less probability.

	Mean	Median	Standard Deviation	MAD	q5	q95	rhat	ess_bulk	ess_tail
Bias Subject A	0.7	0.71	0.07	0.07	0.57	0.82	1.001	6749	5257
Bias Subject B	0.73	0.74	0.1	0.1	0.55	0.88	1.001	6621	5496
Weight1 Subject A	0.59	0.59	0.04	0.04	0.54	0.66	1.0005	4421	2473
Weight 1 Subject B	0.6	0.6	0.04	0.04	0.54	0.67	1.002	5084	2307
Weight2 Subject A	0.58	0.58	0.03	0.04	0.53	0.64	1.0008	4649	2395
Weight2 Subject B	0.56	0.56	0.03	0.03	0.53	0.62	1.0002	3659	2335

Table 2, shows summary outputs for both participants for the parameters in the weighted bayes model.

The models were compared using leave-one-out cross-validation for the model fits. Looking at the values in table x, it became clear that the weighted bayes model best fit the data. This is because the leave-one-out information criteria was lower for the weighted bayes model i.e., 27158.5 compared to 27674.3 for the simple bayes model. However, considering the small sample size for subjects and runs per subject, both model outputs will be entertained in the discussion as no conclusions can be drawn from this proof-of-concept study.

	Simple Bayes (SE)	Weighted Bayes (SE)
ELPD LOO	-13837.2 (56.3)	-13579.2 (75.2)
P LOO	3.0 (0)	23.3 (0.3)
LOOIC	27674.3 (112.6)	27158.5 (150.5)

Table 3, shows model fit estimates for the two models. The estimates are computed using leave-one-out cross-validation. The standard errors for the estimates are written in parentheses.

Discussion

4.1 Differences in Personality and Decision Making

The paper at hand found differences in the decision-making process for the two subjects. In fact, the differences were highly interpretable because decisions were consciously explicit to the subjects themselves, as they assessed the information, they based their decisions on. The subjects were asked to consciously consider the levels of danger associated with opting to ski by using a system created to ostracize the decision from the individual. This conscious effort might explain the extreme similarities between the weight parameters between the subjects. Participants were utilizing a system that weighed the two sources of information for them. However, the system still allowed for individual variability as the *caution, extra caution,* and *not recommended* categories all covered multiple possible combinations of the sources of information. The differences between the two subjects might therefore be explored more efficiently using larger datasets.

There was a difference between the subjects' biases for opting to ski. This might be partly explained by personality differences, as personality traits have been extensively linked with risk-propensity across domains (Mishra & Lalumière, 2011; J. Zhang et al., 2020). Traits such as competitiveness and a preference for a fast-paced life, low levels of anxiety, lack of straightforwardness, lack of self-discipline, and spontaneity have been found to have a positive relationship with risk-propensity (Nicholson et al., 2005). The differences seen in bias towards skiing for the two subjects could be due to differences in one or more of those traits.

To truly capture the full perspective of how different personalities could predict decision making in an extreme sport context, it is necessary to consider the environment of extreme sports. A study on bungee jumpers cognitive ability showed that high arousal in combination with positive valence make people better at certain cognitive tasks (Castellà et al., 2020). They ascribe this phenomenon, partly, to Fredricksons, (2001) *broaden-and-build* theory. This theory proposes that high arousal accompanied with positive valence makes people more attentive and cognitively alert, whereas high arousal and negative valence narrows attention. This theory should be considered for decision-making that happens in ski-touring contexts. Skiers could ski a multitude of different terrain and snow combinations, some of which would naturally invoke higher levels of anxiety in the skier than others. Facing a particularly harrowing run, would therefore potentially make the decision-making of a skier worse than their baseline decision-making.

4.2 Limitations and Considerations for Future Research

Group dynamics might play a role in decision-making in skiing contexts (*Bright, L. S. ,2010*). A paper claimed that individuals who have been part of a group involved in avalanche incidents all report they would have made different decisions had they been on their own (Zweifel, 2015). These findings illustrate a very important shortcoming of the paper at hand. This paper has solely focused on showing differences in risk-propensity from an individual level. Although, when considering that both subjects in the study were expert skiers, their decisions could possibly reflect that of the group's. Their expert status would potentially make them natural leaders of a ski group, and leadership dynamics are highly influential on any group-decision in a mountaineering context (Zweifel & Haegeli, 2014). There is, however, no way of knowing whether the decisions of the subjects would have remained the same, had they skied in different group formations than the ones they potentially did in their recorded samples. Future research would benefit from considering group dynamics.

Future research would also need to solidify the results presented from this study. The study at hand has been purely proof-of-concept and has aimed to show how potential differences in decision-making behavior exist in backcountry skiers even when they consciously decide using the same perceived danger evaluation system. Thus, the author of this paper proposes that Bayesian cognitive models can be helpful tools in exploring these differences. Future studies including ample sample sizes should aim to replicate the findings of this study.

Appendix

Appendix A: Simple Bayes Model

A.1 Prior Posterior Update Checks

Prior posterior update checks were done on the simulated data to assess how well the model learned from the simulated data. It is seen from figure 6 that the model learns from the data and that the model consistently overestimates bias. This is something to be wary about when fitting the model to empirical data.

Figure 6, visualizes prior densities for the simulated skiers with the pink curve. The blue distribution shows the posteriors estimated by the model. The black, dashed line illustrates the naive, simulated mean bias. This is the bias parameter in the simple bayes model.

A.2 Chain mixing

The model estimated its posteriors using two Hamiltonian Markov chains. Visualisations of the two Markov chains show great mixing of the two chains for bias in the simple bayes model and can be seen in figure 7.

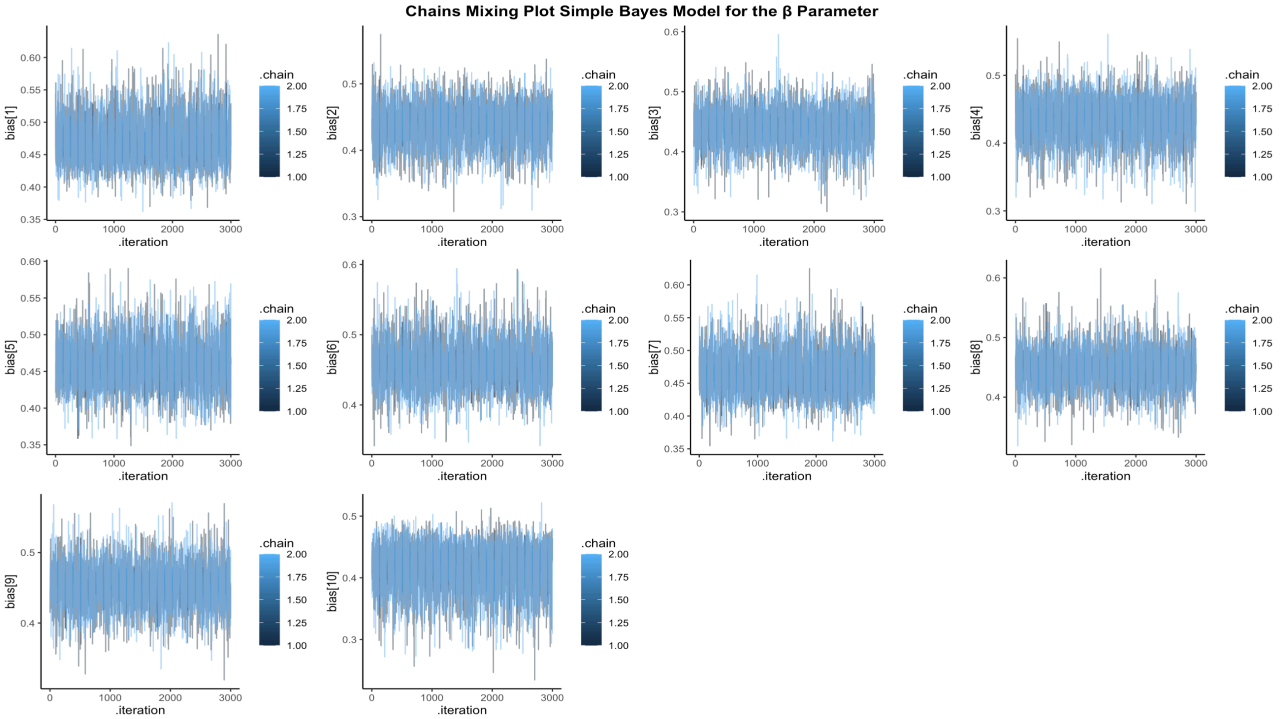

Figure 7, shows that the two Hamiltonian Markov chains specified in the model had great mixing for the bias parameter for all 10 simulated skiers in the simple bayes model.

A.3 Prior & Posterior Predictive Checks

When choosing a prior for the bias-parameter in the model different priors were used in prior predictive checks of the outcome. Thus, 3 meaningful priors for bias were used to assess the influence of prior values on the accumulated outcome of skiing/refraining from skiing. The results can be seen on figure 8.

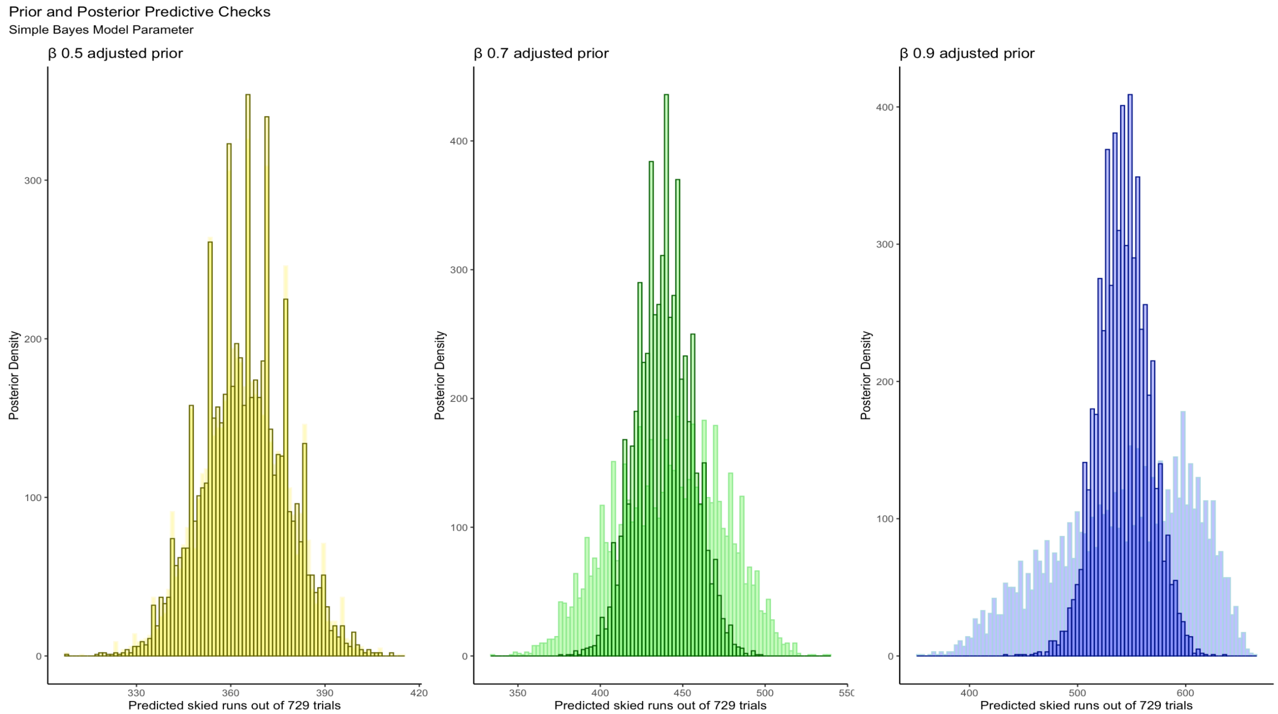

Figure 8, shows prior and posterior predictive checks for the bias parameter using three different priors. The prior used for the predictive checks are displayed in ascending order from left to right. The shaded distributions in lighter colors illustrate the prior predictive checks, while the darker colored overlayed distributions show the prior posterior checks.

It is clear from figure 8 that higher prior values for bias make the distributions narrower for both prior and posterior predicted outcomes. Furthermore, higher values for the prior make the predictions more likely to be that the simulated skiers opted to ski. Looking at this plot the 0.7 adjusted prior is chosen in the model for the bias parameter. This is because this value still allows the model to learn from the data and update the posterior distribution without restricting the outcome too much.

A.4 Prior Sensitivity Analysis

A prior sensitivity analysis was also made to assess how the variability of the priors affected the posterior distribution of the bias-parameter. As seen from figure 9, the prior influenced the posterior distributions as would be expected, i.e., smaller values of standard deviations for the prior distribution resulted in the prior having a larger effect on the posterior distribution.

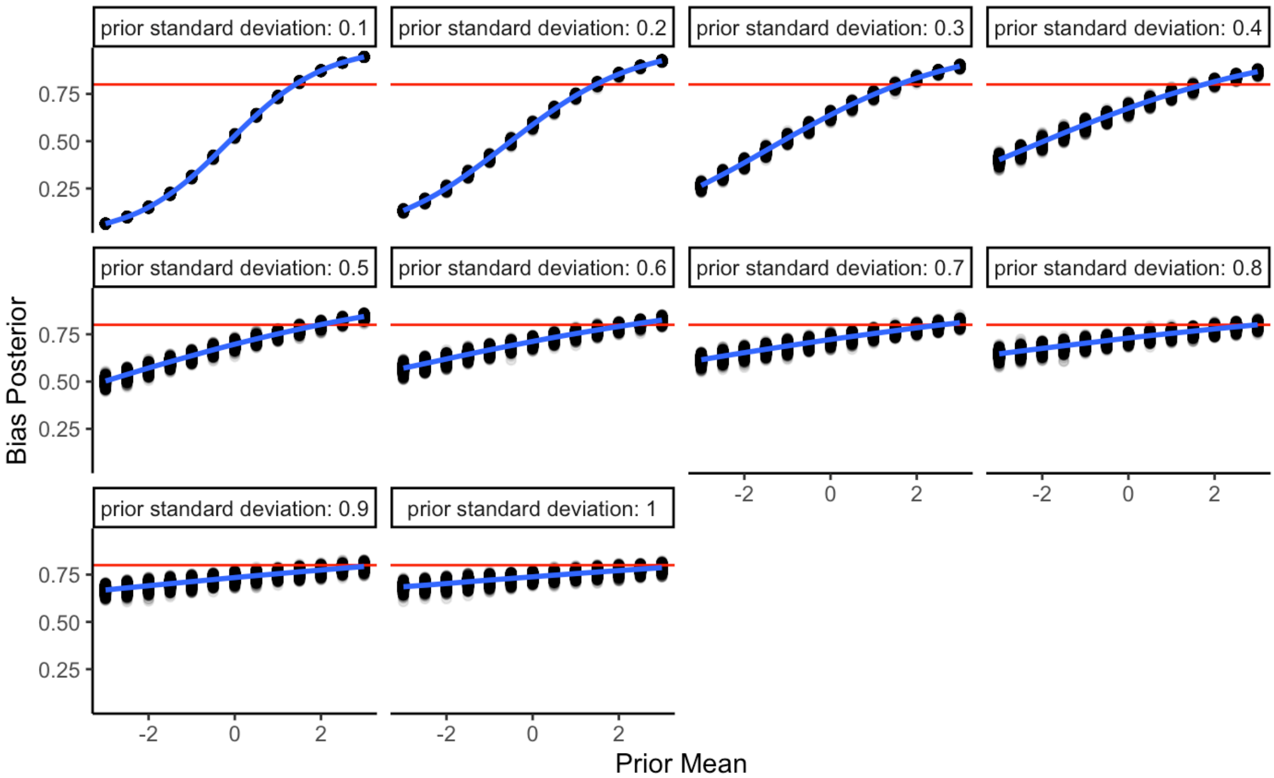

Figure 9, shows how prior values of means and standard deviations influence the posterior distribution of bias.

Appendix A: Weighted Bayes Model

A.5 Prior Posterior Update Checks

Prior posterior update checks were done on the simulated data to assess how well the model learned from the simulated data. It is seen from figure 10 that the model learns from the data and that the model consistently overestimates bias compared to the simulated mean bias. This is something to be wary about when fitting the model to empirical data as any bias value would potentially also be overestimated for the empirical data.

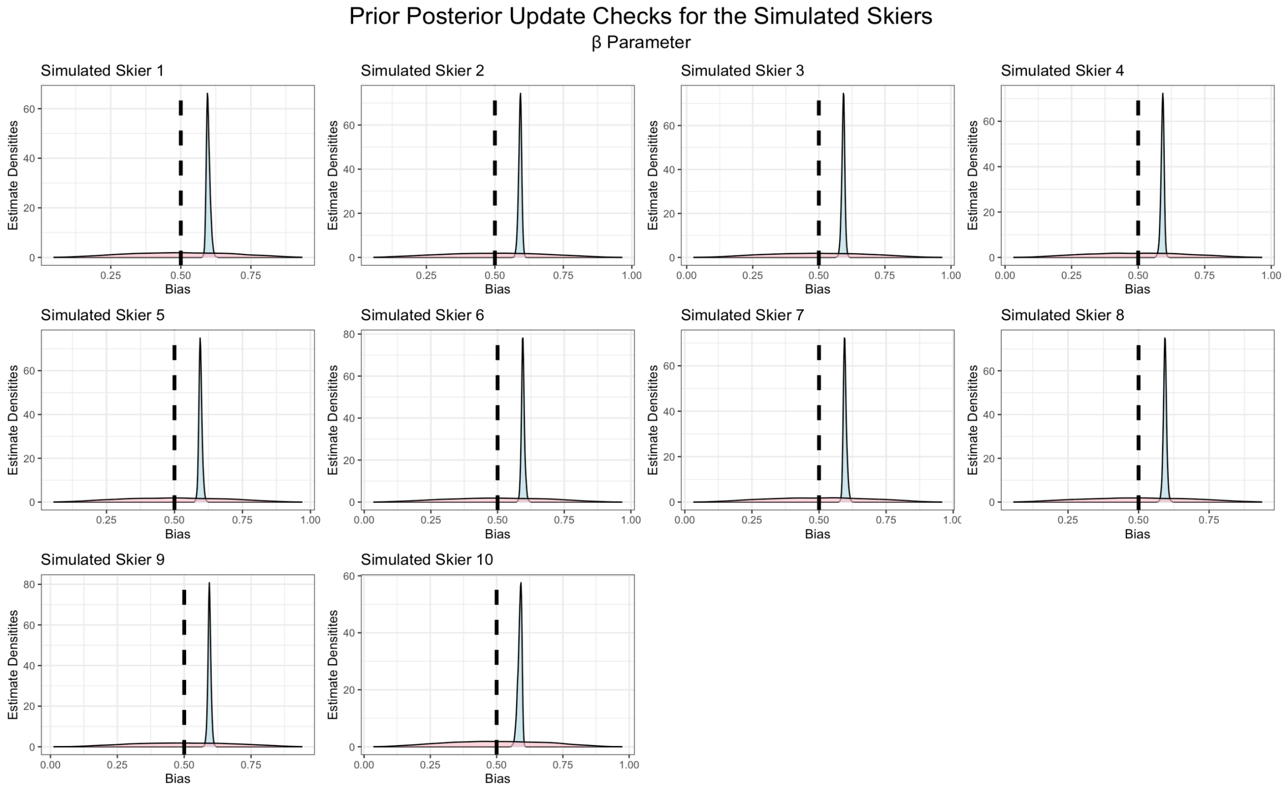

Figure 10, visualizes prior densities for the simulated skiers with the pink curve. The blue distribution shows the posteriors estimated by the model. The black, dashed line illustrates the naive, simulated mean bias. This is the bias parameter in the weighted bayes model.

It is seen from figure 11 that the model updates its posterior estimates of weight 1 from the simulated data and that the model consistently underestimates weight 1 compared to the simulated mean for weight 1. This is something to be wary about when fitting the model to empirical data as any weight 1 value would potentially also be underestimated for the empirical data.

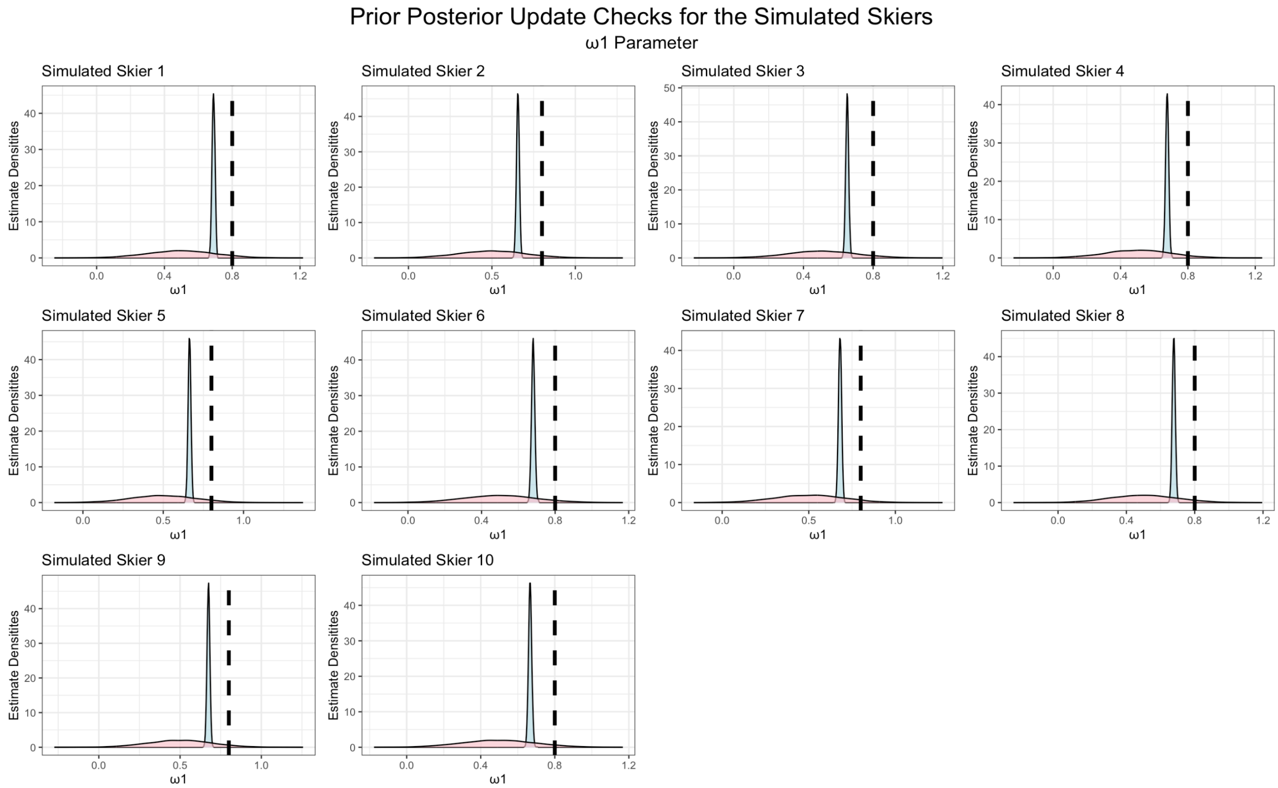

Figure 11, visualizes prior densities for the simulated skiers with the pink curve. The blue distribution shows the posteriors estimated by the model. The black, dashed line illustrates the naive, simulated mean bias. This is the weight1 parameter in the weighted bayes model.

It is seen from figure 12 that the model updates its posterior estimates of weight 2 from the simulated data and that the model estimates weight 2 relatively accurately compared to the simulated mean for weight 2. This means the model could be somewhat trusted to fit an accurate distribution for weight 2 for the empirical data.

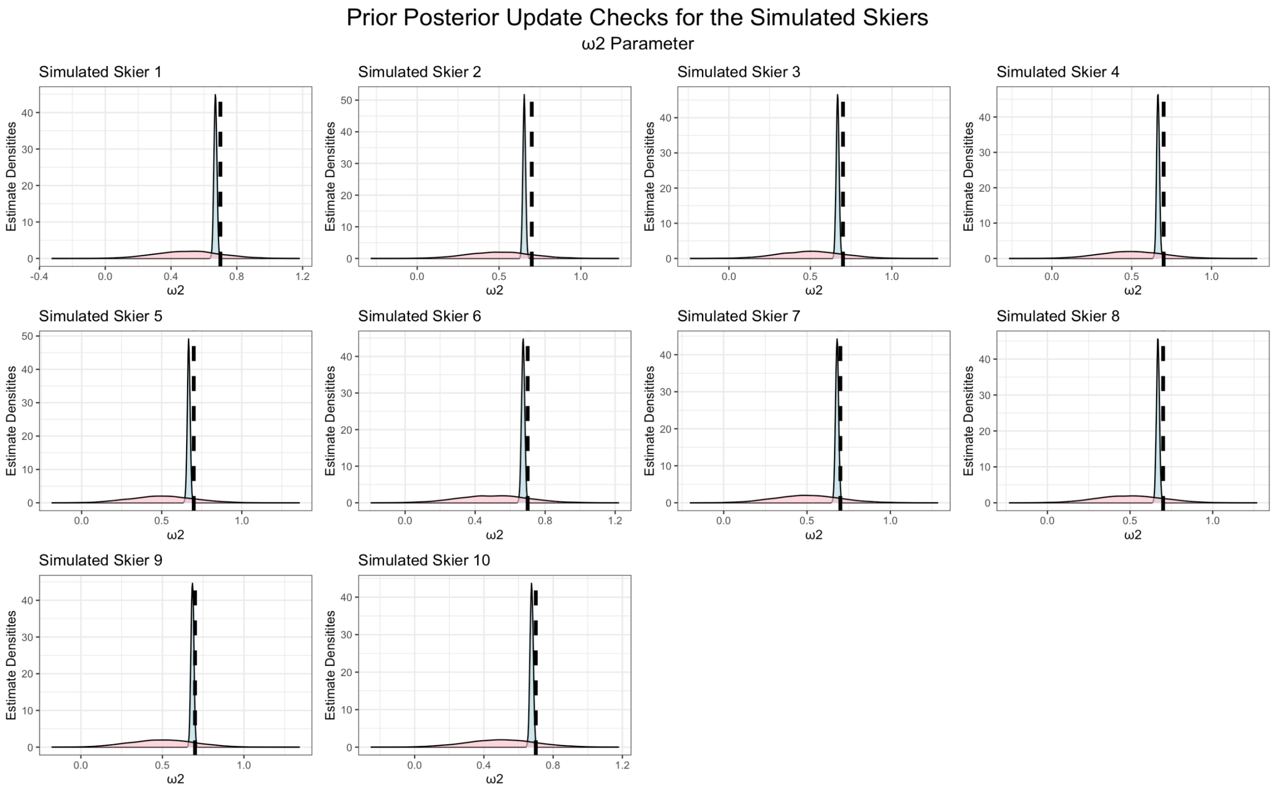

Figure 12, visualizes prior densities for the simulated skiers with the pink curve. The blue distribution shows the posteriors estimated by the model. The black, dashed line illustrates the naive, simulated mean bias. This is the weight2 parameter in the weighted bayes model.

A.6 Chain Mixing

The model estimated its posteriors using two Hamiltonian Markov chains. Visualisations of the two Markov chains show great mixing of the two chains for bias in the simple bayes model and can be seen in figures 13, 14, and 15.

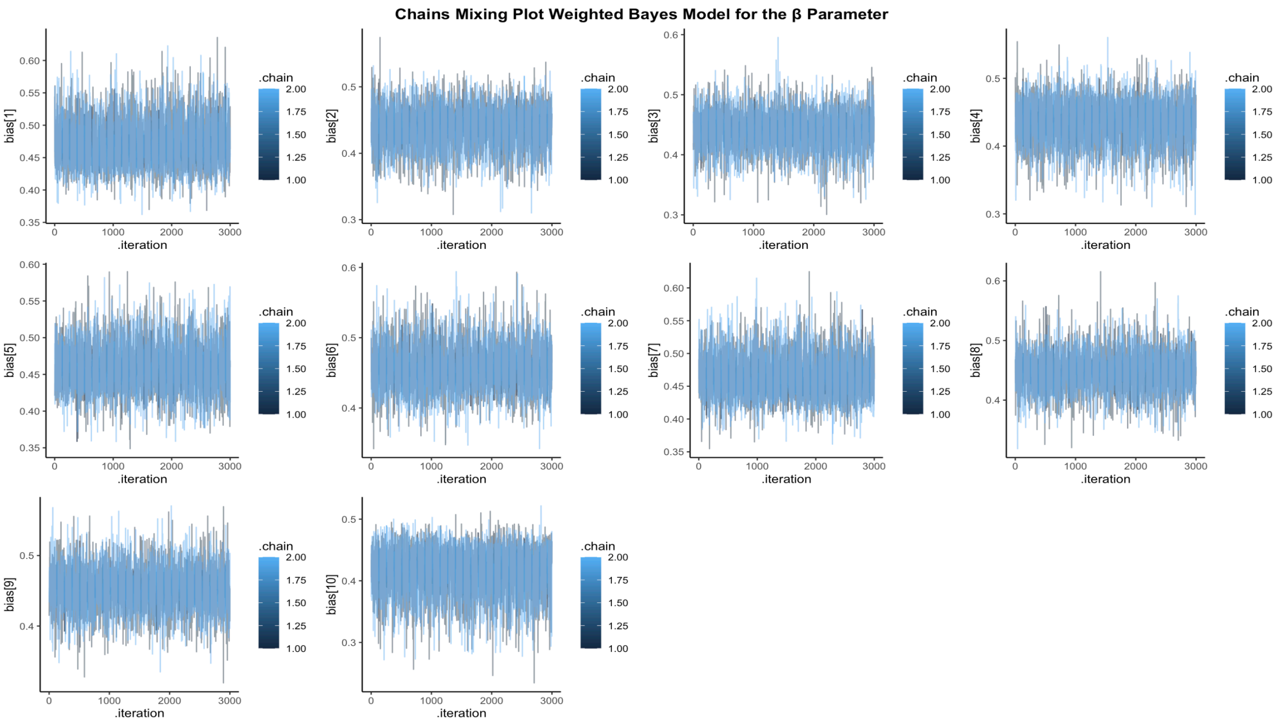

Figure 13, shows that the two Hamiltonian Markov chains specified in the model had great mixing for the bias parameter for all 10 simulated skiers in the weighted bayes model.

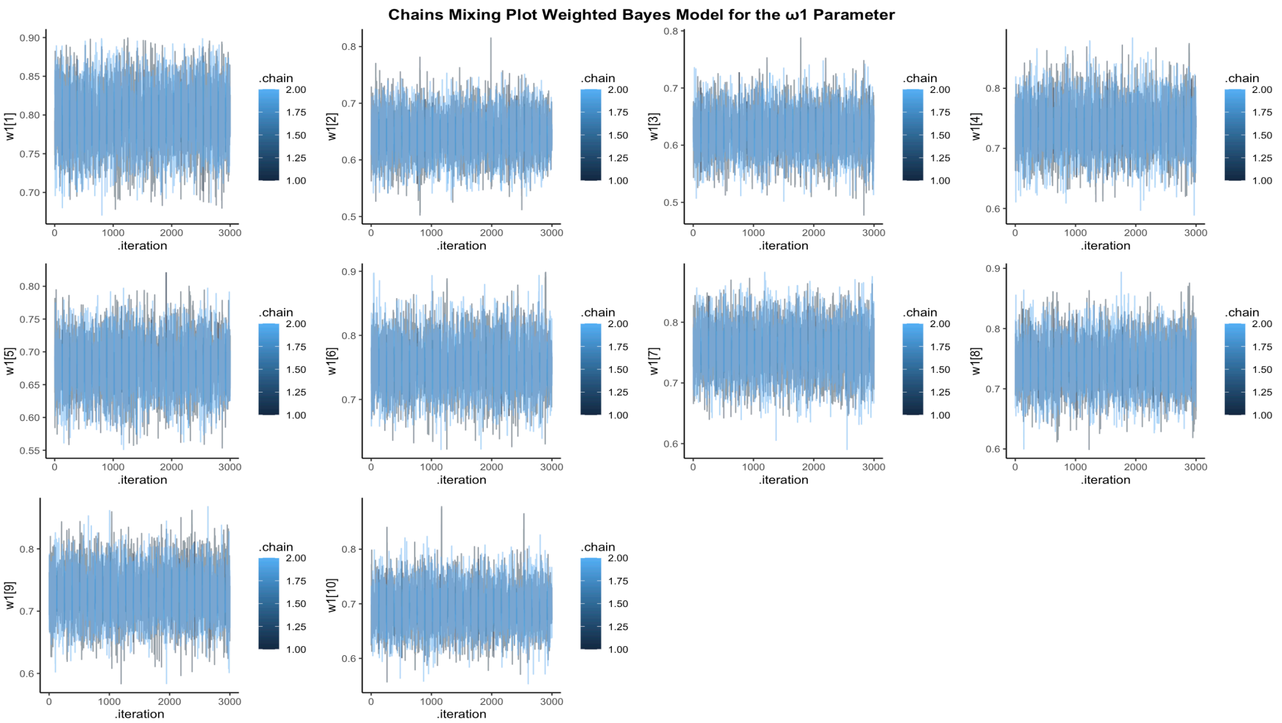

Figure 14, shows that the two Hamiltonian Markov chains specified in the model had great mixing for the weight 1 parameter for all 10 simulated skiers.

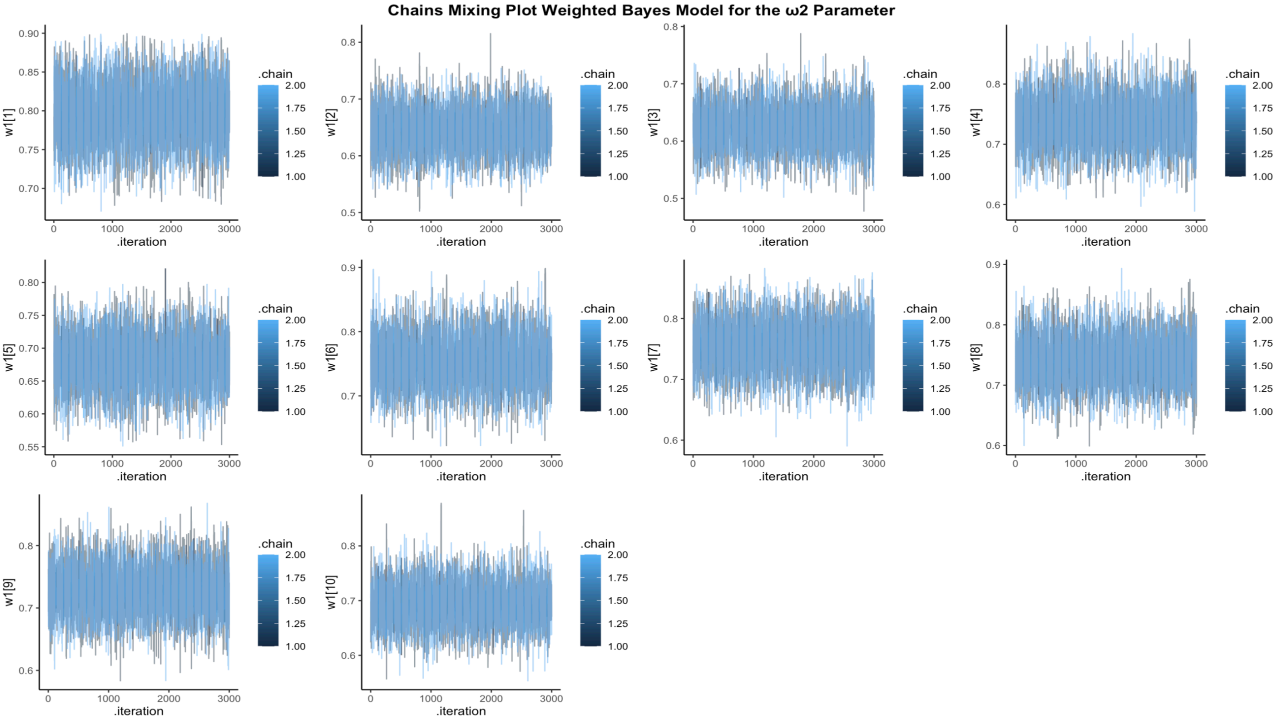

Figure 15, shows that the two Hamiltonian Markov chains specified in the model had great mixing for the weight 2 parameter for all 10 simulated skiers.

A.7 Prior Posterior Predictive Checks

When choosing a prior for the bias-parameter in the model different priors were used in prior predictive checks of the outcome. Thus, 3 meaningful priors for bias were used to assess the influence of prior values on the accumulated outcome of skiing/refraining from skiing. The results can be seen on figure 16.

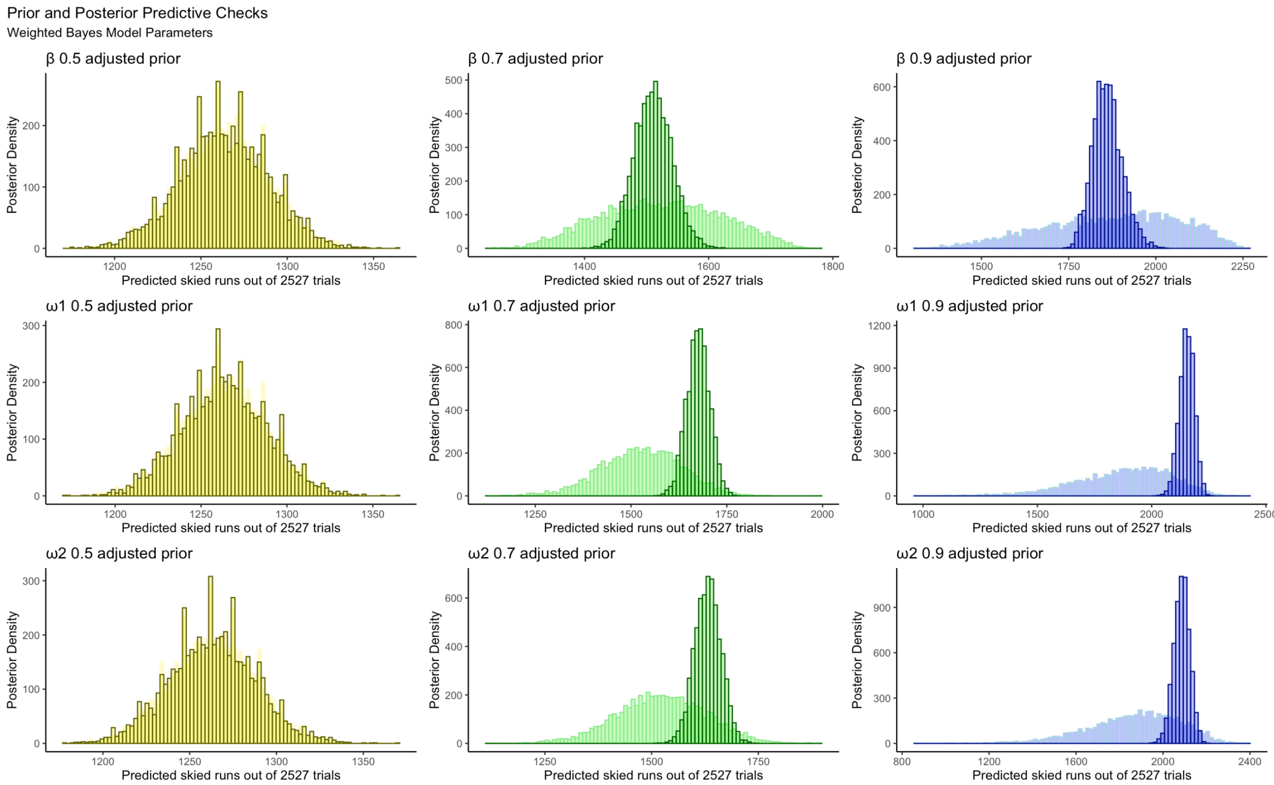

Figure 16, shows prior and posterior predictive checks for the parameters using three different priors. The horizontal grid shows the parameters, while the vertical grid indicates the prior used for the predictive checks in ascending order. The shaded distributions in lighter colors illustrate the prior predictive checks, while the darker colored overlayed distributions show the prior posterior checks.

It is clear from figure 16 that higher prior values make the distributions narrower for both prior and posterior predicted outcomes. Furthermore, higher values for the prior make the predicitons more likely to be that the simulated skiers opted to ski. Looking at this plot the 0.7 adjusted priors are chosen in the model for all parameters. This is because this value still allows the model to learn from the data and update the posterior distributions without restricting the outcomes too much.

B.8 Prior Sensitivity Analysis

A prior sensitivity analysis was also made to assess how the variability of the priors affected the posterior distribution of the parameters in the model. As seen from figures 17, 18, and 19, the priors influenced the posterior distributions as would be expected, i.e., smaller values of standard deviations for the prior distribution resulted in the priors having a larger effect on the posterior distribution for all three parameters in the weighted bayes model.

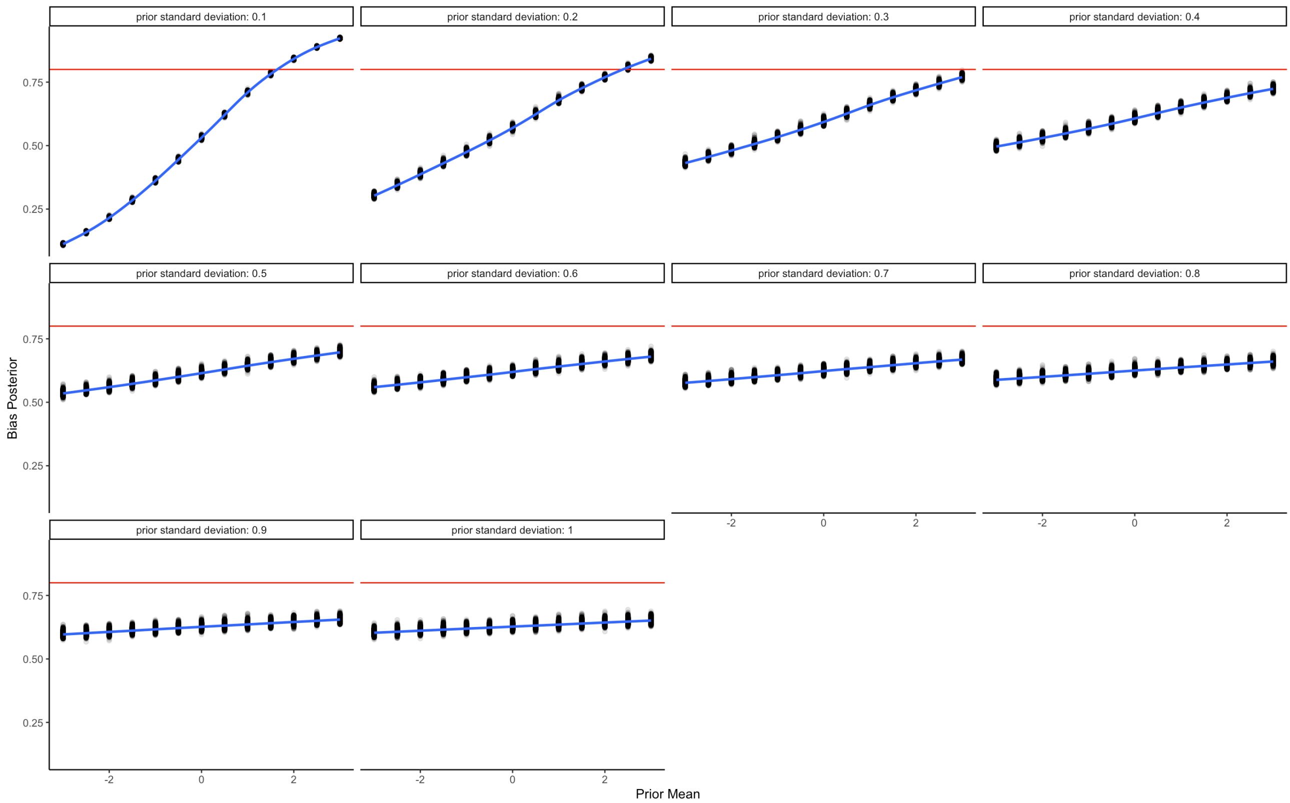

Figure 17, shows how prior values of means and standard deviations influence the posterior distribution of bias.

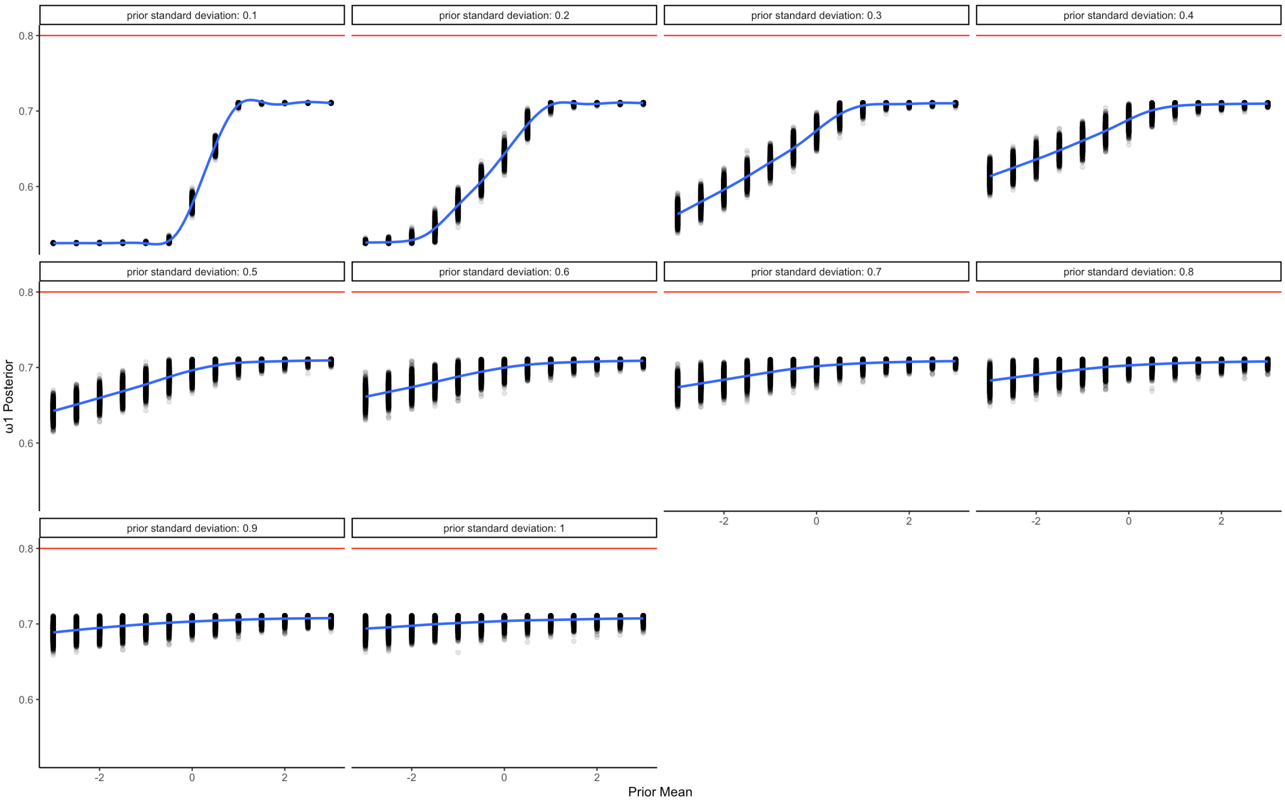

Figure 18, shows how prior values of means and standard deviations influence the posterior distribution of weight 1.

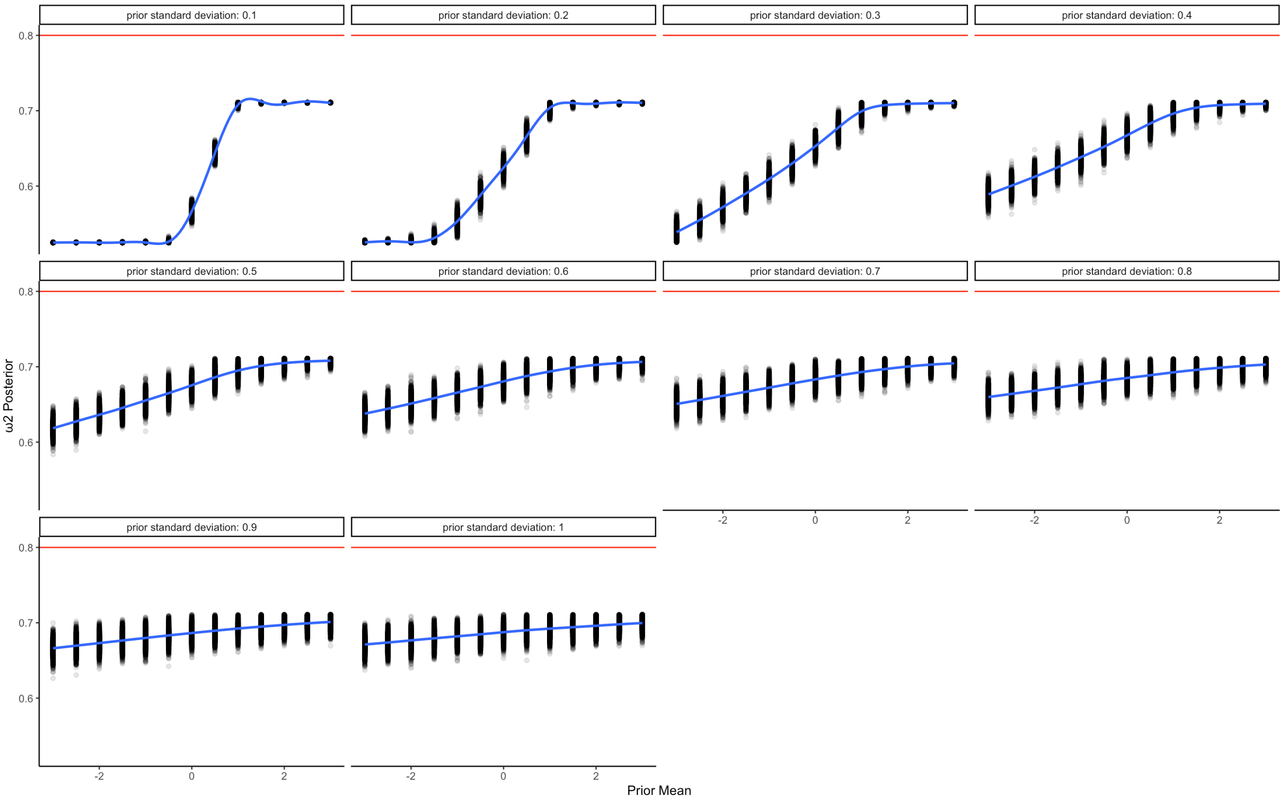

Figure 19, shows how prior values of means and standard deviations influence the posterior distribution of weight 2.