



Norwegian University of
Science and Technology

Modelling and inference of cognitive parameters in jackdaws based on data from an investment problem

Scott Macody Lund

Master of Science in Physics and Mathematics

Submission date: July 2018

Supervisor: Håkon Tjelmeland, IMF

Norwegian University of Science and Technology
Department of Mathematical Sciences

Master Thesis - TMA4905

Modelling and inference of cognitive
parameters in jackdaws based on data from an
investment problem

Scott Macody Lund

Norwegian University of Science and Technology,
Department of Mathematical Sciences,
2018

Preface

This master thesis completes my Master of Science degree at the study programme "Applied Physics and Mathematics" with specialization in "Industrial Mathematics" and further specialization in statistics at the Norwegian University of Science and Technology. The thesis is written at the Department of Mathematical Sciences during the spring of 2018 and is a continuation of my Specialization project (Macody Lund, 2018).

The work is based on an experiment on metacognition in jackdaws performed in 2012 by the neuroscientists Gerit Pfuhl and Robert Biegler at NTNU. I find the experiment very interesting, and it has been fascinating to work with cognitive data from the real world. I am grateful for Gerit and Robert giving me the opportunity to write my master thesis about such an interesting topic. They have also provided me with the data from the experiment, which I thank them for.

My supervisor Professor Håkon Tjelmeland deserves my sincerest gratitude. In addition to finding this very interesting topic, he has given me advice and directions whenever I needed. The work on the original topic metacognition and decision theory have brought up a lot of challenging considerations and discussions. It has been important for me to have someone to discuss with, and Håkon has been a great discussion partner throughout the semester.

Scott Macody Lund
July 12, 2018
Trondheim, Norway

Abstract

Studying animal cognition is both interesting on its own, and because it can improve our understanding of human cognition. It can, for example, reveal when certain cognitive traits were developed in evolution. In this thesis, we consider an investment problem from a behavioural experiment on two jackdaws. The experiment was performed in 2012 by two neuroscientists at NTNU.

In the experiment, the birds are rewarded if they invest enough in the correct situations. The investments are measured by the amount of pecking the birds perform on a computer screen. The birds must evaluate how probable they believe it is that they will be rewarded and how much pecking they want to do. We describe the experiment and develop a statistical model for the situation, including the participants' cognition. Furthermore, we present relevant theory and methods and discuss how they apply to our model. We use these methods to estimate the cognitive parameters and fit our model to the observed data. From the fitted model, new data is simulated and compared to the observed data. Several aspects of the birds' behaviour are found to be present in the simulated data.

There are however some patterns in the birds' behaviour which is not found in the simulations. For one of the birds, a hypothesis test results in the model being rejected at a significance level of 0.05. Based on this, we suggest changes in the model assumptions for that bird. Throughout the analysis, we compare the birds to each other. At one point, we see a surprisingly big difference between them. The data suggest that one bird has much more belief in its own memory when it is hungry. The other bird shows no such effect. Finally, we discuss what could be interesting to consider for further work, including further investigation of whether the birds do respond differently to being hungry.

Sammendrag

Forskning på dyrs kognitive egenskaper er interessant både i seg selv, og fordi det kan hjelpe oss å forstå menneskelig kognisjon bedre. Det kan for eksempel fortelle oss når visse intellektuelle egenskaper ble utviklet i evolusjonen. I denne oppgaven betrakter vi et investeringsproblem fra et atferdseksperiment på to kaier. Forsøket ble utført i 2012 av to nevroforskere ved NTNU.

I eksperimentet blir fuglene belønnet dersom de investerer tilstrekkelig i de riktige situasjonene. Investeringene er målt i hvor mye fuglene hakker på en dataskjerm. Fuglene må vurdere hvor sannsynlig de tror det er at de vil få gevinst og hvor mye de ønsker å hakke. Vi beskriver eksperimentet og lager en statistisk modell for situasjonen, inkludert deltakernes kognisjon. Videre presenterer vi relevant teori og metoder, og diskuterer hvordan de kan anvendes på vår modell. Vi bruker disse metodene til å estimere de kognitive parametrene og til å tilpasse modellen vår til de observerte dataene. Den tilpassede modellen brukes til å simulere nye data, som sammenlignes med observerte data. Det viser seg at flere aspekter av fuglenes adferd er tilstedeværende i de simulerte data.

Imidlertid er det noen observerte adferdsmønstre som ikke viser seg i simuleringene. For en av fuglene resulterer en hypotesetest i at modellen blir forkastet på et 0.05 signifikansnivå. Basert på dette, foreslår vi endringer i modellantagelsene for den fuglen. Gjennom hele analysen sammenligner vi fuglene med hverandre. På et område finner vi en overraskende stor forskjell. Dataene kan tyde på at den ene fuglen har vesentlig større tro på egen hukommelse når den er sulten. Den andre fuglen viser ikke noe tegn til en slik effekt. Tilslutt diskuterer vi hva som kan være interessant å gjøre i videre arbeid, deriblant å undersøke nærmere om det er slik at fuglene responderer ulikt på å være sulten.

Contents

1	Introduction	1
2	Experiment	2
2.1	Experiment description	3
2.2	Experimental data	6
3	Theory	13
3.1	Decision theory	13
3.2	Estimation	16
3.2.1	Maximum performance estimation	17
3.2.2	Maximum likelihood estimation	17
3.3	Bootstrapping	18
4	Model and methods	21
4.1	Model	21
4.2	Decision and utility	23
4.3	Maximum performance estimation	31
4.4	MLE	32
5	Results and discussion	36
5.1	Maximum performance estimation	36
5.2	Maximum likelihood estimation	39
5.3	Simulation	46

6	Closing remarks	55
	Appendices	57
A	Calculation of $E[YI(X \neq Z)I(N = k + 1)r - c Z = z, Y = 0 \cup X = Z \cup N > k]$	57

1 Introduction

One of the most describing features about mankind is our tireless effort to understand the world around us. Science is pushing the frontiers of human knowledge, and we are closer to complete knowledge of the world than ever before. The world around us can be described very precisely and it can be challenging to point out phenomena in daily life which cannot be satisfyingly described. That is until we consider the mind. Concepts such as cognition, intelligence, self-consciousness, free will and others alike, are still far from being understood and arguably holds the biggest unsolved mysteries today. How does our mind work? Why do we act the way we do? How do we make decisions? What makes us conscious, if we are anything other than just randomly interacting atoms at all? Questions like these have befuddled us throughout history and lie at the core of scientific fields such as psychology, neuroscience, philosophy and behavioural analysis. The answers to these questions tell us what governs human behaviour and are very useful in any context involving humans and our decisions. This knowledge is consequentially applicable in many other fields such as economics, politics, sociology, biology (especially evolutionary biology) and psychiatry. For example, when trying to predict the stock market it is relevant to understand the mind of the buyers, and how they are affected by different factors. In psychiatry, a better understanding of the mind can help to detect illness such as schizophrenia. Understanding and describing the different personality types more accurate can improve performance in work life by facilitating the workers better. Information about what makes a person feel satisfied with his or her life situation is interesting in politics, and can also help us live happier lives. Furthermore, these topics may be more relevant than ever, due to the late years' enormous progress in artificial intelligence.

Often questions regarding psychology and behaviour are associated with qualitative theory, in contrast to the quantitative nature of mathematics, physics or chemistry. Historically, psychological knowledge has been argued, discussed and reasoned its way forward, rather than performing quantitative experiments, which is a significant part of the natural sciences. However, as science and technology moved forward we were able to perform more and better experiments which allowed more precise models describing psychology and behaviour. This suggests and enables quantitative theory and the use of precise relations, i.e. equations, also in psychology. In 1961 this was embodied when Richard Herrnstein introduced the matching law (Herrnstein, 1961). He applied mathematical models in the analysis of behaviour and formulated a precise relation between response and reinforcement in a decision

process. This is considered as the founding of the field quantitative analysis of behaviour, a substantial field still today, with its own active community; "The Society for the Quantitative Analyses of Behavior" (SQAB). This use of mathematics allows us not only to understand and describe psychological concepts, but to quantify them precisely. The matching law has later been proven to describe behaviour patterns in numerous areas, such as social dynamics (Borrero et al. 2007), economics and sport. Nonetheless, there is a lot of behaviour where the matching law does not fit.

Quantitative analysis of behaviour is carried out on animals as well as humans, in fact, Herrnstein's paper from 1961 analyzed an experiment on pigeons. This is done both because of interest in animal cognition, but also because it can help us understand humans better. Understanding animal behaviour is interesting for example when trying to answer why we perceive ourselves more conscious than animals. It is also the case that some results on animals may be generalized to humans. For example is the mouse species *Mus musculus* often used in behavioural experiments with the intention of understanding the human brain better, due to its surprising similarities with humans.

Among the people interested in animal cognition are the neuroscientists Gerit Pfuhl and Robert Biegler. They have done research on behaviour and metacognition on birds (Pfuhl & Biegler, 2012) as well as humans (Pfuhl et al. 2012). In 2012 they performed a behaviour experiment, which was an investment problem, on two jackdaws. We have been in contact with them and been given the data from the experiment. The aim of this thesis is to analyze the experiment, and hopefully to reveal some behaviour pattern. Specifically, we will model the investment problem by using the model developed in Macody Lund (2018) and fit it to experimental data. Furthermore, we will suggest a probabilistic model for how the birds make decisions, fit it to the data and discuss the descriptiveness of the model.

2 Experiment

In this section, we first explain the experiment and then present the experimental data. Further analysis will be done after we have introduced the necessary theory and methods.

2.1 Experiment description

The participants in the experiment were two jaybirds, hereinafter referred to as Ari and Susan. The experiment is an investment problem, where Ari and Susan must choose how much effort to put in, in order to be rewarded with a piece of food. The chance of reward is dependent on information that is given to the bird prior to the decision. The birds must, therefore, evaluate this information and take it into account when deciding on an investment. There is also a fixed time interval between the point at which they get the information and when they have to make their decision. Hence they forget some information and have to consider how precise their memories are when choosing the investment. The metacognition of the birds, specifically their belief in their own memory, will hence be reflected in the collected data. Before the experiment was performed the birds did a lot of training so that they know the experiment well. The experiment is a long series of trials spanning over several days. Each trial goes as follows and is illustrated in Figure 1.

- **Step 1:**

The bird is presented with an empty screen with a red dot on it. This is shown until the bird pecks the red dot. When the bird pecks the dot, the trial moves to step 2.

- **Step 2:**

The screen goes blank for a retention interval. With 50% probability, the trial proceeds to 3A and with 50% probability it proceeds to 3B.

- **Step 3A (non-matching case):**

The screen shows a rectangle and a red dot at a different location than the initial dot. The new location can be in any direction from the initial dot, this is uniformly randomly decided. The distance the dot moves, d , can take one of five possible values, this is also uniformly randomly decided. There is now a fixed probability of 20% that there is no chance of being rewarded. If reward is attainable, it is given if the bird pecks the red dot enough times. The necessary number of pecks is a stochastic variable with a geometric distribution with parameter 0.3, resampled for each trial. If the bird is rewarded, the trial ends and after an intertrial interval, a new trial is started by returning to step 1. If the bird at some point pecks on the rectangle in the middle, the trial is ended and after an intertrial interval, the next trial starts.

- **Step 3B (matching case):**

The screen shows a red dot at the same location as the initial red dot

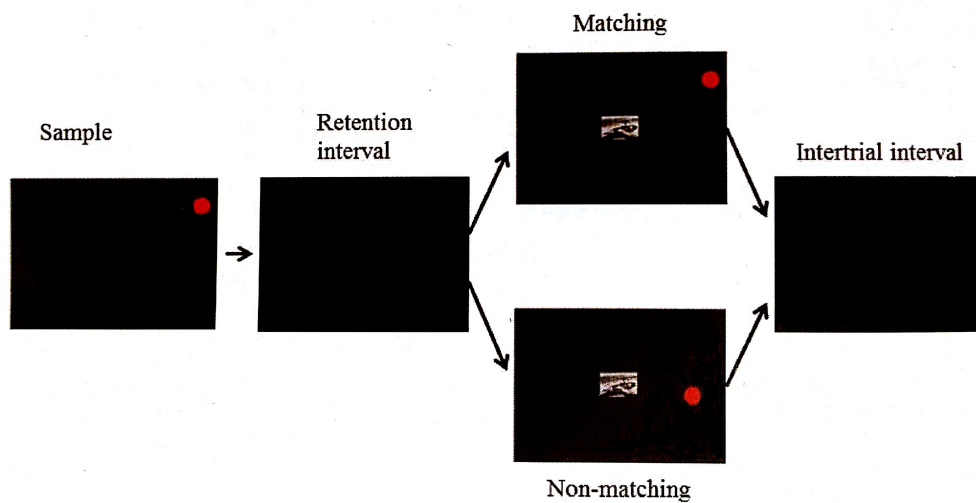


Figure 1: A trial in the experiment. The initial dot is shown till the bird pecks the dot. After a retention interval, it is randomly decided whether the final dot's position is matching or not matching the position of the initial dot. The matching cases are never rewarding, and in these cases the experiment proceeds only when the bird gives up, i.e. pecks the rectangle in the middle. Some fixed proportion of the non-matching cases are rewarding. The reward is attained by pecking enough times, the demanded number is stochastic. The experiment proceeds when the bird is rewarded or if it pecks the rectangle. Non-matching non-rewarding cases behave like matching cases, the experiment proceeds when the bird pecks the rectangle. The next trial starts after an intertrial interval.

and a rectangle. In matching cases there is no possibility of being rewarded. Next trial starts when the bird pecks the rectangle.

We can think of the experiment as a task in detecting whether or not the dot has moved, i.e. if the positions of the initial and final dots differ. In 50% of the trials, the dot has moved. If so, the dot moves one of five possible distances. The possible distances are:

$$d_1 = 1.20, d_2 = 2.15, d_3 = 3.10, d_4 = 4.05, d_5 = 5.00, \quad (1)$$

each distance has the probability $1/5$ of being chosen in the non-matching trials. The general rule is that when the dot has moved, reward is attainable, and the bird should peck. If the dot has not moved, i.e. if $d = d_0 = 0$, the bird should give up and move on to the next trial. However, there are some nuances to this view which we will describe in the following.

The bird is rewarded if and only if it pecks the demanded number of times in a non-matching case with a chance of being rewarded. Hence, if the final dot's position seems to be the same as that of the initial the bird should not invest too much, as there is a high probability that the dot has not moved. If the final dot is far from the bird's memory of the initial dot, it should peck more, since it will be more probable that the dot has moved. The bird must also take into account to what degree it trusts its own spatial memory. If the bird thinks it has an extremely precise memory it should peck a lot even though the dot has just barely moved as long as it is confident that it actually has moved. The biggest investment should be done when the final dot does not match its memory of the initial dot at all and if the bird completely trusts its memory. The investment should still be limited since 20% of the non-matching cases are non-rewarding. In the non-matching cases it will for each peck that does not result in reward become more probable that it is a non-rewarding case, and at some point, the bird should stop no matter how certain it is that the dot has moved. To sum up, the bird should peck more the further away the final dot is from the initial dot because it will feel more certain that the dot has moved. Also, if the bird has already pecked a couple of times without success, it should take this as a hint that it probably is a non-rewarding case, and it should consider stopping pecking.

The information the bird should collect in the experiment, because it is relevant when choosing investment, is the position of the initial dot, and the meta-data which is the proportion of matching cases/non-matching cases, the proportion of non-matching cases that are rewarding, the parameter in the geometric distribution deciding the sufficient number of pecks and how

far the dot moves. The position of the initial dot is specific for that trial and must be memorized in each trial. The meta-data is the same for all trials and the birds will come up with estimates for them during the training period. The experiment is short compared to the training and the metadata in the experiment is identical to that in the training. Therefore it is reasonable to believe that the estimation of these parameters is finished in the training and kept constant during the experiment. The chosen investments, i.e. the amount of pecking, is registered and can be used to evaluate the thoughts of the birds.

2.2 Experimental data

We present the data from the experiment so that we get an idea of the main tendencies in the observed behaviour. The experiment was performed over six consecutive days. At day number three it occurred some technical problems which resulted in no testing of Ari on that day. Hence, we have five days of test data for Ari, a total of 846 trials, and six days of test data for Susan, a total of 1016 trials. Recall that the probability that a trial is a matching case is 0.5 and that the non-matching cases are distributed uniformly over the five distance levels. Ari performed 419 matching trials and about 85 non-matching trials for each distance level. Susan had 510 matching trials and about 100 non-matching trials for each of the five distance levels. For the data presentation that follows, it can be useful to think, as a rule of thumb, that 1% of the matching trials make up roughly 5 trials, while one per cent of the non-matching trials at a certain distance level is about 1 trial.

A natural first question to ask about the behaviour of the birds is how often they gave up without trying to peck, and how this ratio depends on the value of d . This is illustrated in Figure 2. From the figure, we can see that the general trend for both birds is that the more the dot has moved the more often they peck. This pattern is to be expected. Reward is attainable when the dot has moved, and the more the dot has moved the more convinced the bird is that the dot actually has moved, hence it will not give up so often. When $d = 0$, reward is not attainable, and the best thing to do is to give up and move on to the next trial. However, we see in this case that both birds do not give up without pecking in well over half the trials. This indicates that the birds are not too good in realizing when the dot has not moved. Ari has only one percentage point higher rate of giving up without trying when $d = 0$, than when $d = d_1 = 1.20$. Probably he is struggling with noticing

Giving Up Without Pecking

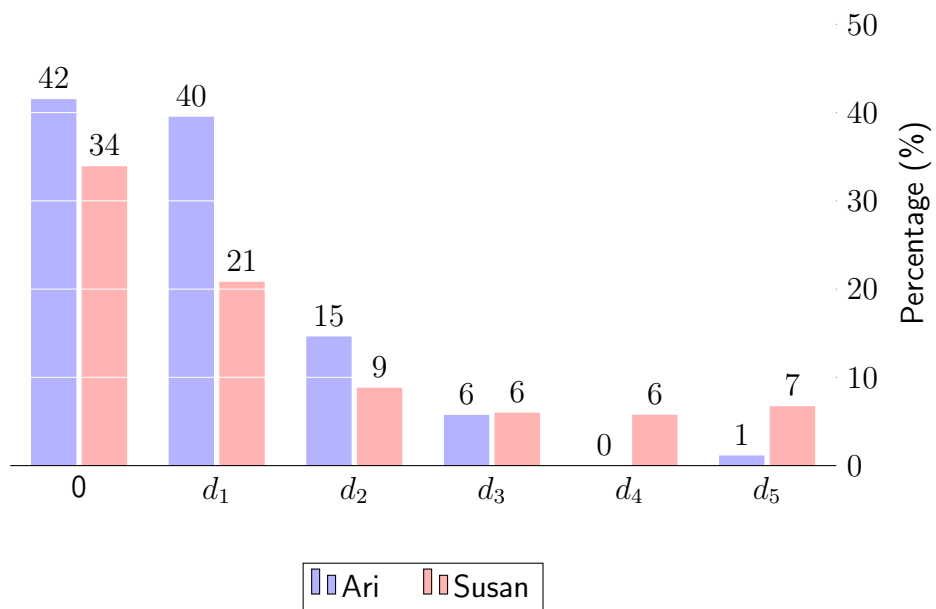


Figure 2: Percentage of trials the birds gave up without pecking. The figure shows a blue bar for Ari and a pink bar for Susan for six different values of d . The matching cases are represented in the leftmost bar-pair where $d = 0$. The non-matching cases make up the five other bar-pairs.

such a small movement of 1.20.

For $d = 0$ the correct response is to give up, we can see that Ari at 42% does a marginally better job at this than Susan at 34%. However, when $d = d_1 = 1.20$ Susan outperforms Ari by only giving up 21% in contrast to his 40%. Moving from $d_1 = 1.20$ to $d_2 = 2.15$ we see a big drop of over 50% in both birds' ratio. It is clear that they are much better at noticing a movement of 2.15 than 1.20. Susan is still doing the best job of having a low giving up-ratio. For larger values of d Susan's ratio flats out at 6 – 7% while Ari's ratio continues to drop all the way to zero, except for one single d_5 -trial where he did give up. He can obviously tell with great confidence when a dot has moved $d_4 = 4.05$ or more. Susan on the other hand also has a low ratio at 6 – 7% for the three biggest distance levels, but she is not as consistent as Ari. She gives up a couple of times even though the dot has moved very far. Comparing the blue and pink bars in the figure we see that the pink bars are evener. Moving left to right, the blue bar starts higher but then declines to a lower level than the pink bars do. Ari's giving up-ratio is varying more with the value of d . This serves him well at $d = 0$, where he correctly gives up more than Susan, and at $d = d_4$ and $d = d_5$ where he correctly does not give up. However, at d_1 and d_2 he incorrectly gives up more than Susan. There is a meeting-point at $d = d_3$ where both birds give up without trying in 6% of the trials. Hence neither of Ari or Susan is clearly doing a better job than the other. Based on this plot, they both have strengths and weaknesses.

The ideal pattern to have in Figure 2 is 100% at $d = 0$ and 0% at all other distance levels. One should have a big decline in the ratio as we see Ari has, but his bars should have been left-shifted. His high giving up-ratio is stretching too far into the high values of d .

We have now seen how the ratio of how often the birds choose to invest, i.e. peck, is for the different distance levels. What about the size of the investment? In addition to more frequent investments for high values of d , we should also find that the investments are bigger. In Figure 3 the average number of pecks is plotted for all distances for both birds.

From Figure 3 there is a clear tendency, the birds peck more for bigger values of d . As discussed earlier we again see that Ari does not invest too much when $d = d_1$, he invests almost one whole peck less than Susan in these cases. Other than that, the two birds have a similar pattern in that they gradually increase their investments with growing value of d . Notice that for both birds the average investment at $d = 0$ is higher than when $d = d_1$. At first glans,

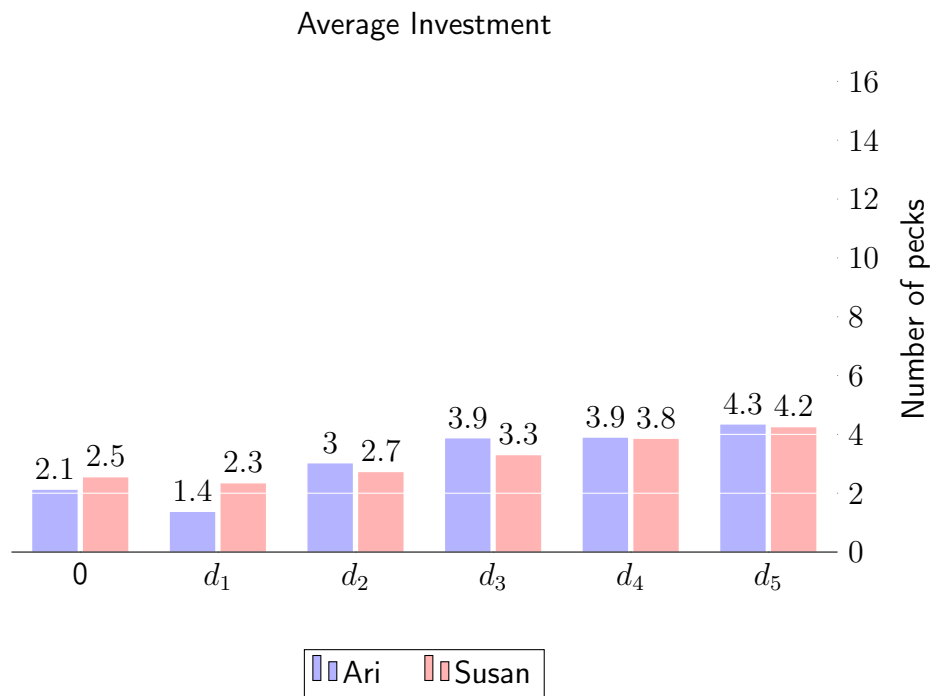


Figure 3: The average number of pecks for each distance level is illustrated, including the trials where the bird gave up, i.e. pecked zero times. The blue bars represent Ari's investments and the pink bars represent Susan's investments.

Average Investment in Non-Rewarding Trials

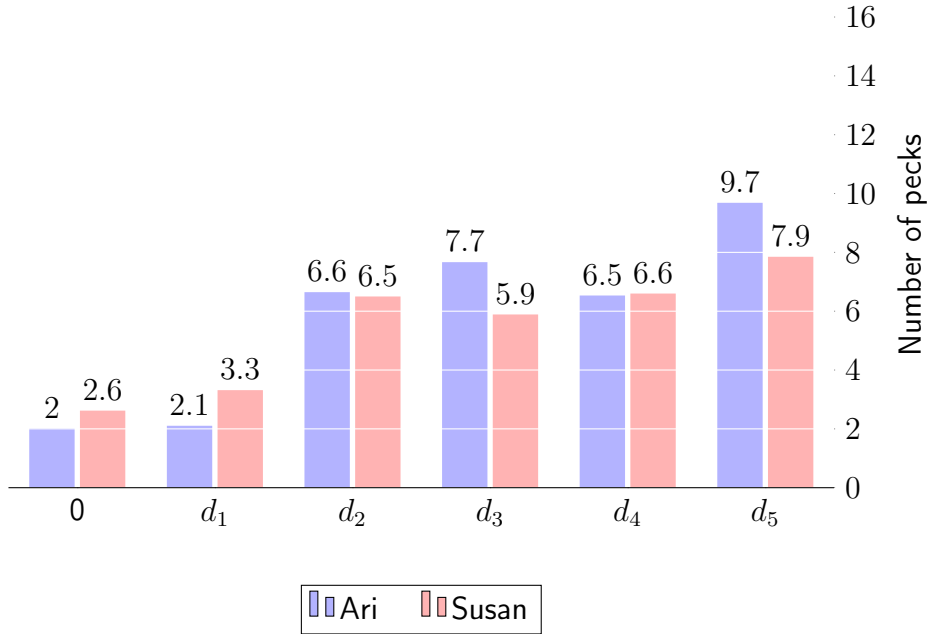


Figure 4: The average number of pecks for each distance level when only considering non-rewarding trials is illustrated. The blue bars represent Ari’s investments and the pink bars represent Susan’s investments.

this might seem unreasonable. However, recall that when $d \neq 0$, 80% of the trials are rewarding. This means that the bird might be rewarded after the first peck, so even if the bird was willing to peck a lot, it was only able to peck once. When $d = 0$, reward is not attainable, hence the bird is never hindered from pecking as much as it wants.

In the 20% of the trials where reward is not attainable the bird is never hindered from pecking. Hence, looking at only these cases we get a picture of how much the birds are willing to peck. In Figure 3 the five bar-pairs for the non-matching cases are held back due to the birds being rewarded. When only using non-rewarding trials we expect the average investments to be higher for non-matching cases ($d \neq 0$) and the same for matching cases ($d = 0$). Furthermore, we expect the investments at d_1 to be higher than the investments at $d = 0$, since the birds should be willing to peck more when the dot has moved 1.20 than when it has not moved at all. The average investment for both birds for all distance levels when only considering non-rewarding trials are plotted in Figure 4.

In Figure 4 we see that investments in non-rewarding trials are bigger at d_1 than at $d = 0$, in contrast to what we saw in Figure 3. The average investment at d_1 was limited by the rewarding trials. Once they are excluded, the bars at $d = 0$ and d_1 are both representing how much the bird is willing to peck, instead of also reflecting their opportunity to peck. When only considering non-rewarding trials, we see an increase in investments at all distance levels, as expected. Note that the bars are not strictly increasing from left to right. This is not alarming as we now are considering pretty small sample sizes. Non-rewarding trials make up only 20% of the trials, and only half of them are non-matching. In addition, that 10% is divided between five distance levels, so only 2% of the trials make up each of the bars where $d \neq 0$ in Figure 4. Consequentially the bars at $d = 0$ is based on roughly 100 trials, while the other is only based on roughly 20 trials. When considering all the data, we can see from Figure 3 that the investments are in fact increasing with the distance d .

We have now looked at a couple of interesting statistics to get an impression of how the birds behaved in the experiment. The next figure is presented to give a picture of the experiment itself, and how often the trials lasted a certain number of pecks. It will also give an overview and summarize some of the observations we have done already. All trials in the experiment end, either because the birds get rewarded or because it gives up. How fast the trials typically end is dependent on the distance. Figure 5 shows the progression in trials with different values of d and gives a picture of how long the different trials lasted. For both birds, we see that the blue graph with $d = 5.0$ is the last one to hit zero. This means that the longest trials, i.e. the trials with most pecks, has $d = 5.0$. This is not surprising as we have already seen that these trials have the largest average investment. We also see that the orange graph with $d = 1.20$ is the first to hit zero for Ari and for Susan it is also the least represented for peck 7-12. This tells us that these trials are quickly ended, even more quickly than when $d = 0$, represented by the red graph. As discussed earlier this is due to the fact that the trials with $d = 1.20$ can be ended both by the bird giving up and being rewarded, whilst the trials with $d = 0$ are only ended by the bird giving up. Especially for Ari, this effect is significant. We can see from Figure 5 that in only 20% of the trials with $d = 1.20$ he reaches the choice of whether to perform the third peck. Another pattern we see is that the six graphs for Ari are more spread than the six graphs for Susan. This supports our observation that, compared to Susan, the behaviour of Ari is differing more between the trials with a low value of d and the trials with a high value of d , made from Figure

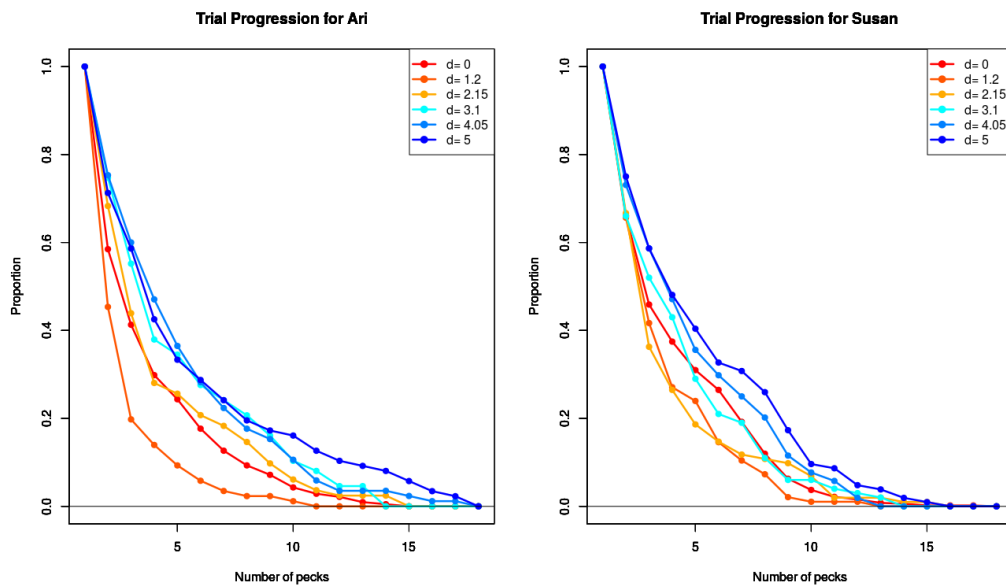


Figure 5: The figure shows how quickly the trials end for different values of d . For some number of pecks, k , on the first axis, the corresponding value on the second axis is the proportion of trials where the bird had the choice of performing the k 'th peck. For $k = 1$ all graphs start at 1.0. This is because, in 100% of the trials, the bird has the choice of performing the first peck since it has not yet had the possibility of giving up or being rewarded.

2. For example, we can see that when $d = 1.20$, Ari almost never reaches peck number 10, but when $d = 5.00$ he does it in about 20% of the trials, while Susan is at peck 10 only ranging from about 0% when $d = 1.20$ to about 10% when $d = 5.00$.

3 Theory

In this section, we present theory that is relevant for the further analysis. We discuss decision theory and introduce some necessary notions within this topic. Two different ways of doing parameter estimation are presented. We also remind the reader of statistical bootstrapping.

3.1 Decision theory

Consider a general choice, for example, what food to order or how to invest your money. Together with a choice, we associate a *possibility set*, A , which is a collection of the possible decisions. In the examples, this could be the set of all dishes served by a restaurant and the set of all stocks in the market. To make a choice we must define a goal, this is embodied mathematically by a *utility function*. The utility measures how good an outcome is. A natural utility function could be how good the food tastes or what the expected return on an investment is, perhaps in combination with the volatility. If there is no randomness related to the choice the utility function simply maps each decision to a utility, that is $U : A \mapsto \mathbb{R}$. In this case, the decision associated with the highest utility should be chosen.

Most often in the real world, there is randomness related to a choice. In that case, the utility will be a function of our decision and some random variables. Suppose the random variables X_1, X_2, \dots, X_n are related to our choice, then $U = U(a, X_1, X_2, \dots, X_n)$, where $a \in A$ is a possible decision. Since the utility is a function of random variables it is itself a random variable. Hence, we can in general not guarantee to receive a satisfying utility from a choice because we can be unlucky. When making a choice related to randomness we instead consider the *expected utility*

$$u(a) = E \left[U(a, X_1, X_2, \dots, X_n) \right]. \quad (2)$$

We define the *optimal decision* to be the decision that maximizes the expected utility,

$$a_{opt} = \arg \max_{a \in A} u(a), \quad (3)$$

and as the name suggests, this is the decision that should be made. To take the expectation in (2) we must know the joint distribution of (X_1, X_2, \dots, X_n) . Sometimes this is given to us by the situation. It is, for example, well known that a fair dice has a uniform distribution with $1/6$ in probability for each side. If this is not the case we must use our experience and knowledge of the system to decide on a *prior distribution* for the variables. By prior distribution we mean the distribution the variable seems to have from our perspective.

Hitherto we have discussed making one choice alone. Often choices are related to other choices, and making some decision might enable us to make more choices. These relations can take many forms, we will in the following consider one specific type of choice. We define a *sequential choice* to be a choice with the natural numbers as its possibility set, that is $A = \{0, 1, 2, \dots\}$, and with the property that in order to make decision k , we must first choose to increase our decision from zero to one, receive the associated utility, increase our decision from one to two, receive the associated utility and so on, all the way up till we increase our decision from $k-1$ to k and receive the associated utility. Thus, such a choice is a sequence of binary choices, each with possibility set $A = \{\text{"increase"}, \text{"not increase"}\}$. For each step in the sequence, we get to know how that decision would have worked out before we have to decide whether or not to increase again. Note that we do not necessarily get to know the realizations of X_1, X_2, \dots, X_n after deciding to increase from zero to one, just the realised utility $U(1, x_1, x_2, \dots, x_n)$.

We give an example of a sequential choice to illustrate what it is and some of the terms we have introduced. Suppose you are reading a book in correct order, and you are looking for the name "NTNU" which you know will appear exactly one time. You have to make the choice, how many pages should you read through before giving up? The possibility set is then the natural numbers as you can choose to read any number of pages. The utility function will be your subjective evaluation of how good or useful it is to find/not find "NTNU" minus the bother it is to read some pages. The true page number of "NTNU" can be considered to be a random variable, which the utility function will depend on. Recall that in order to calculate what decision which gives the highest expected gain we must know the distribution of the involved random variables. If you have almost no knowledge of the book, a uniform distribution where each page is considered having an equal probability of mentioning "NTNU" could be natural to use. If it is a catalogue from NTNU you would probably put almost 100% on page 1. We also see that this choice will satisfy the second criterion for being a sequential choice since once you finish page one, you will know whether it mentioned

"NTNU" or not. You might not know the realization of the random variable, specifically, if you did not see "NTNU" you will not know what page is it on. However you will know how choosing to read only one page would have worked out, i.e. you will receive the utility $U(a = 1, x)$, where x is the true page number.

For sequential choices, we can represent the utility in a slightly different, but useful manner. We define the *expected sequential utility function* evaluated in k to be the expected utility associated with choosing to increase our decision from k to $k + 1$ when already knowing the utility associated with choosing k , that is

$$u_{seq}(k, C) = E \left[U(k + 1, X_1, \dots, X_n) - C \mid U(k, X_1, \dots, X_n) = C \right]. \quad (4)$$

In general, we will have a dependence on C , this embodies the fact that whether or not to increase our decision will depend on any information we have acquired. Hence, this function can not, in general, be evaluated prior to starting increasing our decision. To do that we must perform the first k actions such that we know the value of C . However, this can be simplified if a criterion is satisfied. We will now discuss and identify a necessary criterion such that we remove the dependence on C , such that we can examine u_{seq} without having to consider different cases for C .

In many choices, there is only one element in the possibility set associated with a positive utility for a given realization of X_1, X_2, \dots, X_n . This is the case in our example as the page containing "NTNU" will be the only page with positive utility. This will be the case in most situations related to searching for some object and deciding for how long to look. In these situations, we would, of course, stop looking as soon as we find what we are looking for. In terms of the expected sequential utility function, finding the correct page would make u_{seq} negative for all the following pages. Hence, u_{seq} would be dependent on C . Now, we restrict ourselves to cases of choices with only one element in the possibility set associated with a positive utility for a given realization of the random variables and when this decision has not yet been done. In such cases, nothing else is revealed to us during the choice than whether or not we have found what we are looking for. Hence, when only considering cases prior to the finding, when it is still useful to look more, we have no C -dependence. This is because there is only one possible information we can get after reading k pages and that it is still useful to look more, namely that we did not find "NTNU" on the first k pages. Since we knew that this information would be present if we came to this situation,

we can include it in u_{seq} from the start. This is done by putting C to be whatever cost there is of reading the first k pages. Our criterion for being able to represent the expected sequential utility function without C is consequentially a searching-choice where we by searching learn nothing other than whether we have succeeded or not. Note that there will technically be a dependence on C as we by succeeding our search will make all further searching useless. However, we are now only interested in the cases where the search has not ended. We obviously would stop searching once succeeded. We can, therefore, meaningfully, for these choices set

$$u_{seq}(k) = E \left[U(k+1, X_1, \dots, X_n) - C(k) \mid U(k, X_1, \dots, X_n) = C(k) \right], \quad (5)$$

where $C(k)$ is the cost of reading k pages. With this notation, $u_{seq}(k)$ is the expected utility from choosing to read page $k+1$ when already having read pages $1, 2, \dots, k$ without finding "NTNU". Summarized, we eliminated C by restricting us to choices where its only possible value is the cost of searching.

The expected sequential utility function represents the expected utility of the sequence of binary choices of whether or not to increase the decision by 1. It is tempting to say that if $u_{seq}(k)$ is positive we should increase, and if it is negative we should give up and not increase. However, in sequential choices, we should be willing to make some choices with a negative expected utility if it allows us to later make choices with high enough expected utility. We call this implied utility. For example, if you knew page 100 mentioned "NTNU" you could be willing to read page 99 knowing that it would not give positive utility since you then later would be able to read page 100 and get positive utility. Again, we can simplify by adding a criterion. We require that $u_{seq}(k)$ is non-increasing as a function of k . Under this requirement, we can use the decision rule to increase when $u_{seq}(k)$ is positive and not increase when $u_{seq}(k)$ is negative. There will be no implied utility since once $u_{seq}(k)$ becomes negative, it will not become positive later.

3.2 Estimation

Given an experiment which is some choice, we can calculate the expected utility of making some decision. In behavioural analysis, the utility will typically depend on some cognitive parameters, θ , of the participant. In our example earlier, a reader will experience some bother from reading a page, the magnitude of that bother is a parameter. Thus, the expected utility

function can be written as

$$u = u(a, \boldsymbol{\theta}) = E \left[U(a, X_1, \dots, X_n) \right], \quad (6)$$

where a is a decision, and the cognitive parameters $\boldsymbol{\theta}$ are incorporated in both the function U and in the distribution of X_1, \dots, X_n . Suppose we want to estimate these parameters and we had some participant perform the experiment. The participant made some decision $a_k \in A$. How should we use this information to make statistical inference about $\boldsymbol{\theta}$? This is a complicated question of psychological and philosophical nature as well as statistical. There is no clear answer, but we will discuss two possible methods in the following.

3.2.1 Maximum performance estimation

We define the *performance*, Q , of a participant to be the chosen expected utility divided by the optimal expected utility, that is

$$Q(\boldsymbol{\theta}) = \frac{u(a_c, \boldsymbol{\theta})}{u(a_{opt}(\boldsymbol{\theta}), \boldsymbol{\theta})}, \quad (7)$$

where a_c is the chosen decision of the participant and $a_{opt}(\boldsymbol{\theta})$ is the optimal decision given the parameters $\boldsymbol{\theta}$. The performance is a measure of how good decisions a participant makes. With optimal behaviour we have $Q = 1$. We can then define the *maximum performance estimator* to be

$$\hat{\boldsymbol{\theta}}_{\text{MPE}} = \arg \max_{(\boldsymbol{\theta})} Q(\boldsymbol{\theta}). \quad (8)$$

The estimate is the parameter values that maximize the performance of the participant. Another interpretation is that we tune the parameters such that the quality of the choices is as high as possible.

3.2.2 Maximum likelihood estimation

Maximum likelihood estimation is a fundamental estimation method in statistics. Our presentation of the topic is based on Warpole et. al. (2012). Likelihood is a measure of the plausibility of a model given some observed data. In maximum likelihood estimation (MLE) we estimate parameters to have values such that the probability to sample the observed data from the model is as high as possible. Suppose the random variables $Y_1 \dots, Y_n$ are independent

and identically distributed according to $f(y|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is unknown. Let the realizations $Y_1 = y_1, \dots, Y_n = y_n$ be given. Now the question is, under what value of $\boldsymbol{\theta}$ was this data generated? In MLE, we estimate $\boldsymbol{\theta}$ to be the value such that the probability of sampling the observed data, called the likelihood $L(\boldsymbol{\theta}|y_1, \dots, y_n)$, is as high as possible, that is

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}|y_1, \dots, y_n) = \arg \max_{\boldsymbol{\theta} \in \Theta} \prod_{i=1}^n f(y_i|\boldsymbol{\theta}). \quad (9)$$

Differentiating a product of many factors is a much more difficult task than differentiating a sum of many terms. Recall that $(uv)' = u'v + uv'$, while $(u + v)' = u' + v'$. To make this maximization easier we define the log-likelihood $l(\boldsymbol{\theta}|y_1, \dots, y_n) = \log(L(\boldsymbol{\theta}|y_1, \dots, y_n))$. The product in (9) will then become a sum as $\log(ab) = \log(a) + \log(b)$, and the problem is much simplified. The logarithm is a bijection, hence the ordering is reserved. This implies that the value of $\boldsymbol{\theta}$ which maximizes the log-likelihood will also maximize the likelihood. Hence we maximize the log-likelihood instead, for example by differentiating and solving

$$\frac{\partial l(\boldsymbol{\theta}|y_1, \dots, y_n)}{\partial \boldsymbol{\theta}} = 0. \quad (10)$$

In order to perform MLE we need a probability distribution, i.e. $f(y|\boldsymbol{\theta})$. Note that this is not necessary to perform MPE, where we can simply use the utility function. In MLE we maximize a probability, so in order to perform it we need a distribution describing the probability that some decision are made,

$$P(a_c = a) = f(a|\boldsymbol{\theta}). \quad (11)$$

We can then perform MLE by estimating $\boldsymbol{\theta}$ to the value maximizing the probability of making the chosen decision a . This probability should depend only on the expected utility. The reason for this is the flexible nature of the concept of utility. Utility is simply whatever the participant wants to maximize. So any effect influencing the probability of some decision can be embedded in the utility.

3.3 Bootstrapping

We now present statistical theory on bootstrapping, much of the following presentation is inspired by Givens & Hoeting (2013). Bootstrapping is a statistical method where resampling of observed data is used to make inference. It is typically used when we do not know anything about a distribution,

except for some observed sample from that distribution. The sample will, therefore, contain all available information, and the best way to further investigate the distribution is to resample the observed sample.

Suppose we are given a sample of some data $\mathbf{x} = \{x_1, \dots, x_n\}$ from some generating mechanism. The generating mechanism can be described by some probability distribution F which is unknown to us. Let us assume that all we know is that all the sample points x_1, \dots, x_n come i.i.d. from the same unknown distribution F . We define the *empirical distribution*, \hat{F} , to be the distribution that puts a probability of $1/n$ at each of the observations x_1, \dots, x_n . Heuristically one can think that we use our sample to construct a distribution that imitates F . It should be noted that sampling from \hat{F} is equivalent to drawing from the original sample $\mathbf{x} = \{x_1, \dots, x_n\}$ with replacement.

Sampling from \hat{F} , we can produce arbitrarily many, for example, B , new samples of the same size, called *bootstrap samples*. We denote the bootstrap samples $\mathbf{x}^{*(1)}, \dots, \mathbf{x}^{*(B)}$. All elements of all bootstrap samples will be i.i.d from the empirical distribution, that is

$$x_1^{*(b)}, x_2^{*(b)}, \dots, x_n^{*(b)} \stackrel{iid}{\sim} \hat{F} \quad b = 1, 2, \dots, B. \quad (12)$$

Suppose we are interested in a parameter of the distribution, $\theta = t(F)$. One way to estimate θ is to use the so called *plug-in principle*, that is $\hat{\theta} = t(\hat{F})$. Hence if we have some way to compute the parameter given the distribution, the plug-in principle is to use the empirical distribution instead of the true distribution.

The estimation of the parameter θ based on data \mathbf{x} is done using some *estimator*, i.e. $\hat{\theta} = s(\mathbf{x})$. Hence, if we have an estimator s we can from our original sample make an estimate of θ . But what is the associated uncertainty to this estimate? We want to estimate $SD_F[\hat{\theta}]$, i.e. the standard deviation of the estimate of θ when sampling from F . We now use the plug-in principle, and replace the true distribution with the empirical distribution. Hence

$$\widehat{SD_F[\hat{\theta}]} = SD_{\hat{F}}[\hat{\theta}^*], \quad (13)$$

where $\hat{\theta}^* = s(\mathbf{x}^*)$ is an estimate of θ based on a bootstrap sample \mathbf{x}^* . This estimator is called the *ideal bootstrap estimator*. Since \hat{F} is known, we have

all the necessary information to compute the estimate in (13). Therefore, we can in theory compute the ideal bootstrap estimate, and this is to prefer as we then get the lowest possible uncertainty. However this is not always easily done. Instead we can use each of our B bootstrap samples to compute one estimate for θ , i.e.

$$\widehat{\theta}^{*(b)} = s(\mathbf{x}^{*(b)}) \quad b = 1, 2, \dots, B. \quad (14)$$

The empirical standard deviation, or standard error, of these B estimates will then be an estimate of the ideal bootstrap estimator. This estimate is called the bootstrap estimate of the standard deviation of the estimator $\widehat{\theta} = s(\mathbf{x})$, i.e.

$$\widehat{SD}_F(\widehat{\theta})_{bootstrap} = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\widehat{\theta}^{*(b)} - \widehat{\theta}^{*(\cdot)})^2}, \quad (15)$$

where $\widehat{\theta}^{*(\cdot)} = \frac{1}{B} \sum_{b=1}^B \widehat{\theta}^{*(b)}$. This latter method is very simple, but it will give more uncertainty than the ideal bootstrap estimator. By performing the sampling of the B bootstrap samples we are adding an error called *Monte Carlo error*, hence making the estimate more uncertain.

Suppose we have estimated the parameter θ for both birds, and found that $\widehat{\theta}_{\text{Ari}} \neq \widehat{\theta}_{\text{Susan}}$. When comparing the birds to each other, a natural question to ask is whether or not the difference, $\widehat{\delta}_\theta = \widehat{\theta}_{\text{Ari}} - \widehat{\theta}_{\text{Susan}}$, is statistically significant. Do we have sufficient evidence to claim that there is a difference between the birds? To answer this, we can use a hypothesis test. Let $H_0 : \delta = 0$ be that there is no difference between the birds, and let $H_1 : \delta \neq 0$ be that there is a difference. A natural test statistic to use is $T = \delta$. As before, when using a complicated estimator, an analytical solution can be difficult to find. In stead we can use bootstrapping to explore the distribution of the difference $\widehat{\delta}_\theta = \widehat{\theta}_{\text{Ari}} - \widehat{\theta}_{\text{Susan}}$. Under H_0 we can mix the data from Ari and Susan since there is no difference between them. From the mixed data we sample for Ari B bootstrap samples of size 846, as that is the size of his data set. Likewise we sample B bootstrap samples of size 1016 from the mixed data for Susan. The reason we must use their respective sample size is that the estimated difference based on the original data was generated in that way. Hence we are interested in the distribution of the difference when using 846 trials for Ari and 1016 for Susan. We can then compute B estimates of δ by

$$\widehat{\delta}_\theta^{(b)} = \widehat{\theta}_{\text{Ari}}^{(b)} - \widehat{\theta}_{\text{Susan}}^{(b)}, \quad \text{for } b = 1, \dots, B. \quad (16)$$

The distribution made up by $\widehat{\delta}^{(b)}$, called the bootstrap distribution, is an estimate of the distribution of $\widehat{\delta}_\theta$ under H_0 . We now consider the estimate, $\widehat{\delta}_\theta = \widehat{\theta}_{Ari} - \widehat{\theta}_{Susan}$, we get from using only Ari’s data for his estimate and Susan’s data for her estimate. If there is no difference between Ari and Susan, this estimate is a realisation from the bootstrap distribution. If the estimate is far out in the tail of the bootstrap distribution, it is unreasonable that it comes from that distribution. In such a case, we reject the null hypothesis. Specifically, at a 95% significance level, we reject H_0 if $\widehat{\delta}$ is in the lower or upper 2.5% of the bootstrap distribution.

When performing bootstrapping we need to decide on a value of B , that is how many bootstrap samples we shall draw. The value depends naturally on what we are estimating. As a general rule of thumb, $B = 200$ should be sufficient when estimating an average or the standard deviation. If we want to estimate the behaviour in the tail we need a bigger B , since there are less data in the tail. When we use the middle 95% of the bootstrap distribution above a rule of thumb is that $B = 1000$ suffice. However, is the bootstrap sampling algorithm runs quick, there is nothing that hinders us from using an even bigger value of B . The estimates will only become better with increasing B .

4 Model and methods

In this section, we go through how the theory is applied to our experiment and introduce more specifically the methods we will use. We will also present a model for the experiment.

4.1 Model

We now present the model we will use when analyzing the experiment quantitatively. This model was developed in Macody Lund (2018), where the same experiment was analyzed. Recall the procedure of the experiment described in Section 2. The experiment starts with a blank screen with a red dot on it being presented to the bird. This screen is the domain for which the entire experiment will take place, we denote it \mathcal{D} . Furthermore, we denote the initial red dot’s position by $\mu \in \mathcal{D}$. At this moment in time, the value of μ is precisely known to the bird. Once the bird is ready it pecks the red dot, the dot now disappears, and the screen goes blank for a retention interval.

During the retention interval, the bird forgets the exact position of the initial

red dot. However, it has a memory of the position. To have a memory of the position can be thought of as, having some idea of how probable it is that different locations on the screen, in fact, was the true position of the initial dot. Our model is that the bird considers the position of the initial dot to be a random variable X . A normal distribution centred at the true position of the dot seems reasonable, hence

$$X \sim N(\mu, \sigma^2), \quad (17)$$

where σ^2 is a parameter descriptive of the spatial memory of the bird.

After the retention interval, the experiment is either in stage 3A or 3B, this is unknown to the bird. Either way, a final dot Z is shown on the screen. Its position matches that of the initial dot in 50% of the cases, i.e. if it is stage 3B. In the non-matching cases, the generating mechanism described in Section 2 was used to decide the position of the final dot. It is unreasonable to believe that the bird understands this mechanism. Two features of the mechanism it might notice is that the distribution of the final dot is point symmetric about the initial dot and that the area close to the initial dot has the highest probability density. We choose to model this as a normal distribution centred at the initial dot, which also has these two features. Hence the distribution of the final dot's position is

$$f(z|x) = \frac{1}{2}\delta_x(z) + \frac{1}{2}n(z; x, \tau^2), \quad (18)$$

where $\delta_x(z)$ is the dirac delta function which puts a probability of 1 at x and $n(z; x, \tau^2)$ is the density at z for a normal distribution with expectation x and variance τ^2 . The distance between the initial dot and the final dot is what we refer to as the distance level $d = |z - \mu|$.

Recall that some fixed proportion of the non-matching trials are not rewarding. This information is hidden in the experiment in the sense that the bird can not know whether a non-matching trial is rewarding or not. Hence it must consider this to be stochastic. Let Y be a Bernoulli distributed variable with success probability p_Y ,

$$f_Y(y) = \begin{cases} 1 - p_Y, & \text{for } Y = 0, \\ p_Y, & \text{for } Y = 1, \end{cases} \quad (19)$$

and let $Y = 0$ indicate that a trial is not rewarding and $Y = 1$ indicate that it is rewarding. In the experiment, $p_Y = 0.8$ was used, meaning that for a non-matching trial it is 4 times as probable that it is rewarding than that it is not.

Symbol	Description
\mathcal{D}	Screen
X, μ, σ^2	Initial dot
Z, τ^2	Final dot
Y, p_Y	Rewarding/not reawarding
N, p_N	Necessary pecks

Figure 6: Symbols describing the experiment.

Let N be the number of necessary pecks in order to be rewarded, this is geometrically distributed, hence we set

$$f_N(n) = (1 - p_N)^{n-1} p_N, \quad (20)$$

where p_N is the success probability. The value in the experiment was $p_N = 0.3$, so the expected number of necessary pecks is $E[N] = 1/0.3 = \frac{10}{3}$. The established notation is summarized in Figure 6.

4.2 Decision and utility

Consider the experiment to be a choice where we must decide how many times to peck. Recall that the bird is hindered from pecking more if rewarded, so strictly speaking the choice is how many times the bird is willing to peck. The possibility set is then

$$A = \{0, 1, 2, \dots\}. \quad (21)$$

In order to make a decision, we should identify a utility function. We must hence try to imagine what the bird is trying to achieve, what does it want? Food is something all animals have an instinctive urge for as this is highly necessary to survive. Since the reward is a piece of food, there should be a positive utility associated with getting the reward. It is not the case that an animal always is in a constant search for food. The effort to find food is dependent on many variables including time since last meal and activity level. However, in this project, we will for simplicity assume a constant utility $r > 0$ of being rewarded.

Performing labour is a limited resource which must be spent wisely. It takes both focus and uses up calories, hence we associate a cost $c > 0$ with having

to perform a peck. Also here we assume that all pecks have the same cost, implying that to perform n pecks have n times as high cost as a single peck. This assumption is not indisputable but nonetheless used in this thesis. The reward and the peck are the only things we will directly associate a utility with. The following analysis has to do with the probability of being rewarded for a certain number of pecks in different situations, this is to calculate the expected utility.

We now establish the dependence the utility has on the random variables X, Z, Y and N and the parameters r and c . What does it take to be rewarded? First of all the dot has to have moved, that is we need $X \neq Z$. Furthermore, recall that some of the non-matching cases were not rewarding. For the rewarding cases, we have that $Y = 1$. Hence, if $X \neq Z$ and $Y = 1$ reward is attainable. Reward is given if the bird pecks enough. That is, if the number of pecks, k , it is willing to perform is greater than N , i.e. $k > N$. We can then conclude that the bird is rewarded if and only if

$$YI(X \neq Z)I(k > N) = 1, \quad (22)$$

where $I(B)$ is the indicator function, which is equal to 1 if B is true and 0 otherwise.

In a rewarding non-matching trial the birds pecks $\min\{k, N\}$ times, otherwise it pecks k times. Each peck cost c , hence the utility function is

$$\begin{aligned} U(k, Y, X, Z, N) &= YI(X \neq Z)I(k > N)r \\ &\quad - YI(X \neq Z)\min\{k, N\}c - (1 - YI(X \neq Z))kc. \end{aligned} \quad (23)$$

Oftentimes the next step would be to compute the expected utility and encourage to make the decision that maximizes the expected utility. However, recall that the possibility set of this choice is $A = \{0, 1, 2, \dots\}$ and that the pecks are performed in a sequence. This is thus a sequential choice. The experiment can be viewed as a long sequence of binary choices where the bird must choose whether or not to peck. The possibility set for each choice in the sequence is

$$A = \{\text{peck, not peck}\}. \quad (24)$$

As discussed in Section 3 we can now define the expected sequential utility $u_{seq}(k)$. This is the expected utility associated with increasing our number of pecks from k to $k + 1$ when knowing that the first k pecks did not result in reward. If we choose to peck once more, reward is attained if and only if $Y = 1$, $X \neq Z$ and $N = k + 1$. We condition on the fact that the first k

pecks are unrewarded, that is, $Y = 0 \cup X = Z \cup N > k$, and of course on the position of the final dot. The expected sequential utility is hence

$$\begin{aligned} u_{seq}(k|Z = z) = \\ E[U(Y, X, Z, N, k + 1) - kc|Z = z, Y = 0 \cup X = Z \cup N > k] = \\ E[YI(X \neq Z)I(N = k + 1)r - c|Z = z, Y = 0 \cup X = Z \cup N > k]. \end{aligned} \quad (25)$$

The calculation of this expectation is complicated, hence we leave it to appendix. The result is

$$\begin{aligned} u_{seq}(k|Z = z) = \\ \frac{p_Y \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2)(1 - p_N)^k p_N \cdot r}{(1 - p_Y) + p_Y \mathcal{N}_z^\mu(\sigma^2, \sigma^2 + \tau^2) + p_Y \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2)(1 - p_N)^k} - c, \end{aligned} \quad (26)$$

where

$$\mathcal{N}_{\mathbf{x}_0}^\mu(\sigma^2 + \tau^2, \sigma^2) = \frac{n(\mathbf{x}_0; \boldsymbol{\mu}, \sigma^2 + \tau^2)}{n(\mathbf{x}_0; \boldsymbol{\mu}, \sigma^2) + n(\mathbf{x}_0; \boldsymbol{\mu}, \sigma^2 + \tau^2)}. \quad (27)$$

It is also worth mentioning that this latter expression has a interesting interpretation. The probability, from the birds' perspective, that the dot has moved, when knowing the position of the final dot, $Z = z$, is

$$P(X \neq Z|Z = z) = \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2). \quad (28)$$

This expression is hence the probability that a trial is non-matching, this is also shown in Appendix. The sequential utility depends on the parameters p_y , p_N , σ^2 , τ^2 , r and c , however not all of them are equally interesting to estimate. Additionally, estimation in higher dimensions is a challenging task, so we should avoid estimating too many parameters. The parameters p_N , p_Y and τ^2 were decided by the experimenter and kept constant throughout the experiment. One could however still consider these to be cognitive parameters and estimate what the bird's perception of these parameters are. On the other hand, the true values of these parameters are known, in contrast to that of σ^2 , r and c which are true cognitive parameters. We know nothing about their values as they exist only inside the brain of the bird. Consequentially, with limited computational power, we choose to not estimate p_N , p_Y nor τ^2 . We rather consider these to be constant and have the values used in the experiment. The values for the two former were $p_N = 0.3$ and $p_Y = 0.8$. Regarding the latter, recall that the actual generating mechanism used was not a normal distribution which we have used in our model. That means the experimenter did not explicitly choose a value for τ^2 . We must hence use another method of assigning a value to this parameter. Given that the bird operates with a normal distribution for the position of the final dot in the

non-matching cases, it is reasonable to use MLE to estimate τ^2 based on the observations of Z . For the non-matching cases, our model states that the final dot has a binormal distribution

$$f(z_1, z_2) = \frac{1}{2\pi\tau^2} \exp \left\{ -\frac{1}{2} \left(\frac{z_1^2 + z_2^2}{\tau^2} \right) \right\}, \quad (29)$$

where z_1 and z_2 are the two coordinates of the final dot, and τ^2 is the variance in each coordinate. We are interested in the distribution as a function of d , hence we use that $d^2 = z_1^2 + z_2^2$ and integrate over all point equidistant from the origin

$$f(d) = \int_0^{2\pi} \frac{1}{2\pi\tau^2} \exp \left\{ \frac{-d^2}{2\tau^2} \right\} d \cdot d\theta = \frac{d}{\tau^2} \exp \left\{ \frac{-d^2}{2\tau^2} \right\}. \quad (30)$$

All distance levels are used equally often. We therefore assume that the bird has observed five realizations of d , one at each distance level d_1, \dots, d_5 . The likelihood becomes

$$L(\tau^2 | d_1, \dots, d_5) = \prod_{i=1}^5 \frac{d_i}{\tau^2} \exp \left\{ \frac{-d_i^2}{2\tau^2} \right\}, \quad (31)$$

and we get the following log-likelihood

$$l(\tau^2 | d_1, \dots, d_n) = \log L(\tau^2 | d_1, \dots, d_n) = \sum_{i=1}^5 \log(d_i) - \log(\tau^2) - \frac{d_i^2}{2\tau^2}. \quad (32)$$

To maximize the log-likelihood we solve

$$\frac{\partial l(\tau^2 | d_1, \dots, d_n)}{\partial \tau^2} = \sum_{i=1}^5 \frac{d_i^2}{2\tau^4} - \frac{1}{\tau^2} = 0. \quad (33)$$

Multiplying the latter equation with $-\tau^2$, dividing by 5 and moving the terms with d_i to the other side we get

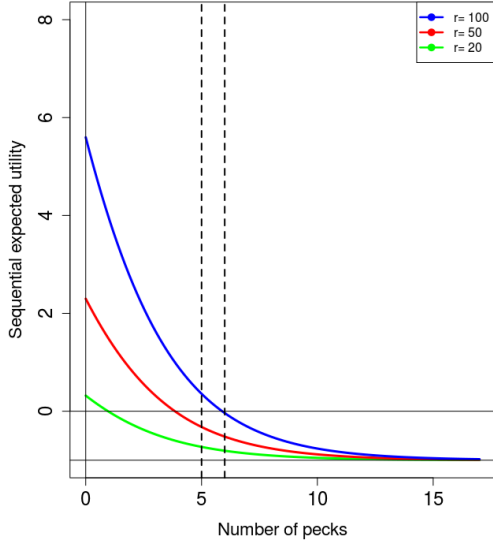
$$\widehat{\tau^2}_{\text{SME}} = \frac{1}{10} \sum_{i=1}^5 d_i^2 = 5.7075. \quad (34)$$

As with $p_N = 0.3$ and $p_Y = 0.8$, we now also consider $\tau^2 = 5.7$ to be a constant which we do not estimate.

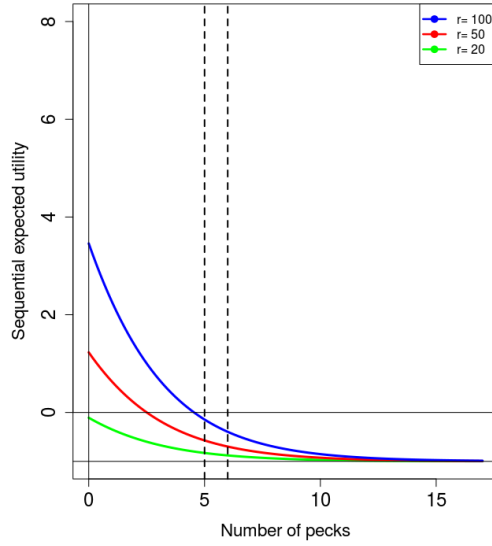
The parameters r and c are the only two which is a measure of utility. Since

there is no standard unit for utility, we can without loss of generality let $c = 1$. Thus, a peck costs one unit of utility, and the value of r will represent how many pecks the reward is worth. We conclude that the two parameters of interest are σ^2 and r . Thus, we investigate how the sequential utility depend on these parameters. In Figure 7 the sequential expected utility is plotted for different values of σ^2 and r , for the other parameters we have used $p_N = 0.3$, $p_Y = 0.8$, $\tau^2 = 5.7$ and consider the distance level $d_1 = 1.20$. From Figure 7 we see that the blue curve is always above the red curve which is always above the green curve. This means that for every binary choice of whether or not to peck, a higher value of r gives a higher expected utility by choosing to peck. Also, we see that the bigger value of r the later the curve hits zero. Recall that pecking should be done as long as the sequential utility is positive, hence the figure shows that for all three values of σ^2 more pecking should be performed for higher values of r . Note that strictly speaking the sequential expected utility is only defined for the natural numbers $0, 1, \dots$, however, we can see from (26) that $u_{seq}(k|Z = z)$ is a continuous function of k which can be interesting to graph. To decide whether or not to peck we must consider the graphs in Figure 7 evaluated at the integers. We see that when $\sigma^2 = 0.2$ the blue graph is above and the other two are below zero. This means that if $r = 100$ and the bird has already pecked five times, the bird should peck once more, while if $r = 50$ or $r = 20$ it should not peck. Once six pecks have been performed, all three graphs are below zero and pecking should not be done. By similar observations we conclude that if $\sigma^2 = 1$ and $r = 100$, pecking should not be done when $k = 5$ nor when $k = 6$. For σ^2 and $r = 100$ the utility by pecking another time when already having done six pecks is positive. Hence we should be willing to perform the seventh peck in this case.

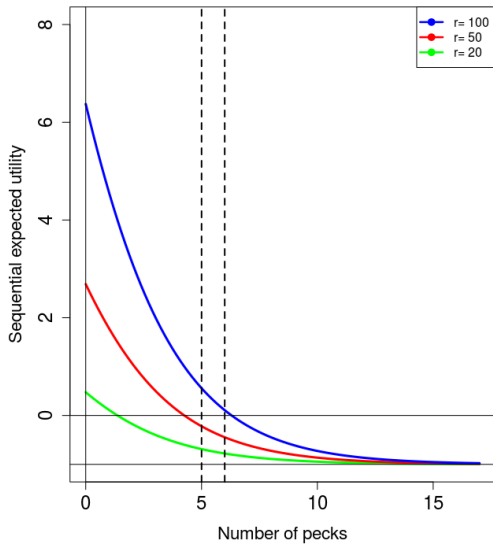
The effect r has on the utility is both straightforwardly and probably as expected. The bigger value of r , the more valuable reward is, and therefore being rewarded will have a higher utility and more pecking should be done in order to achieve it. On the other hand, the effect σ^2 has is slightly more intricate. Figure 7 shows that the utility for $\sigma^2 = 0.2$ and $\sigma^2 = 4$ are more similar to each other than they are to the utility for $\sigma^2 = 1$. To understand this, we revisit the situation the bird is presented with. The bird is shown a final dot at a distance $d = 1.20$ from the initial dot and what we model to be the centre of the probability distribution the bird puts on the initial dot. Then the question is, is the final dot the same as the initial dot, or is it a different dot? Or formulated differently, does this final dot come from a distribution with variance σ^2 or $\sigma^2 + \tau^2$? The bird experiences a distance of 1.20, but is this due to a bad memory or due to a movement and a bad memory? If the distance between the final dot and the centre of its memory is



(a) $\sigma^2 = 0.2$



(b) $\sigma^2 = 1$



(c) $\sigma^2 = 4$

Figure 7: Sequential expected utility for different values of σ^2 and r . Dashed lines are drawn at 5 and 6. Other parameters are kept at their respective constant value and $d = 1.20$.

1.20 when $\sigma^2 = 0.2$, it is very unlikely that this movement is only due to the variance in the memory. Such a precise memory would not be that mistaken. The dot is very likely to have moved, i.e. been drawn from a distribution with variance $\sigma^2 + \tau^2$, and hence there is a high utility of pecking because reward is probably attainable. When increasing to $\sigma^2 = 1$ the density of the distribution representing the bird's memory becomes higher at $d = 1$. This means that it becomes more probable that the dot has not moved and that the perceived movement of $d = 1$ is only perceived and not real. Hence the bird should not peck as much. Increasing the forgetfulness of the bird further to $\sigma^2 = 4$ the density representing the memory evaluated in $d = 1$ decreases due to spreading. If the dot has moved, it is from the birds perspective drawn from a distribution with variance $\sigma^2 + \tau^2$, this distribution will also spread when we increase σ^2 , but not as much as the former will due to the constant τ^2 . Recall that we are considering which of these two distributions the final dot come from. Even though the distribution associated with a movement spreads, it becomes more probable compared to the alternative. From (27) and (28) we can see the ratio of interest. Hence the further increase of σ^2 when $d = 1$ makes it seem more probable to the bird that the dot has moved and more pecking should be done.

Each of the binary choices of whether or not to peck is characterized completely by the distance level d and the number of already performed pecks k . For matching cases, $d = 0$, and for non-matching cases, d takes one of five values. In total there are six different values for d . Both Ari and Susan performed at most 16 pecks in a trial, so the value of k is ranging from 0 to 16, that is 17 different values. Hence there are no more than $6 \cdot 17 = 102$ different binary choices in the experiment. All of these 102 choices did not necessarily appear in the experiment. For example, did neither of the birds peck 16 times for $d = 0$, so the choice with $d = 0$ and $k = 16$ did not appear. However, both birds did over 800 trials, most choices appeared several times. Each of these 102 choices is associated with an expected utility by pecking. For a higher value of d , it will seem more probable that the dot has moved and the utility will be higher. The bigger value of k , the more probable it is that either it is a matching case or a non-matching non-rewarding case so the utility associated with continuing to peck will be lower. The sequential expected utility in the 102 choices for $r = 50$ and $\sigma^2 = 1$ is illustrated in Figure 8. We see that pecking two times is correct even though $d = 0$. This means that when $d = 0$ it will not be so obvious that the dot has moved, and if being rewarded is worth the cost of 50 pecks it is, in fact, correct to peck 2 times. The black line is for $d = 5$ between $k = 11$ and $k = 12$. Hence, if 11 pecks have been performed, the expected utility for pecking another time is

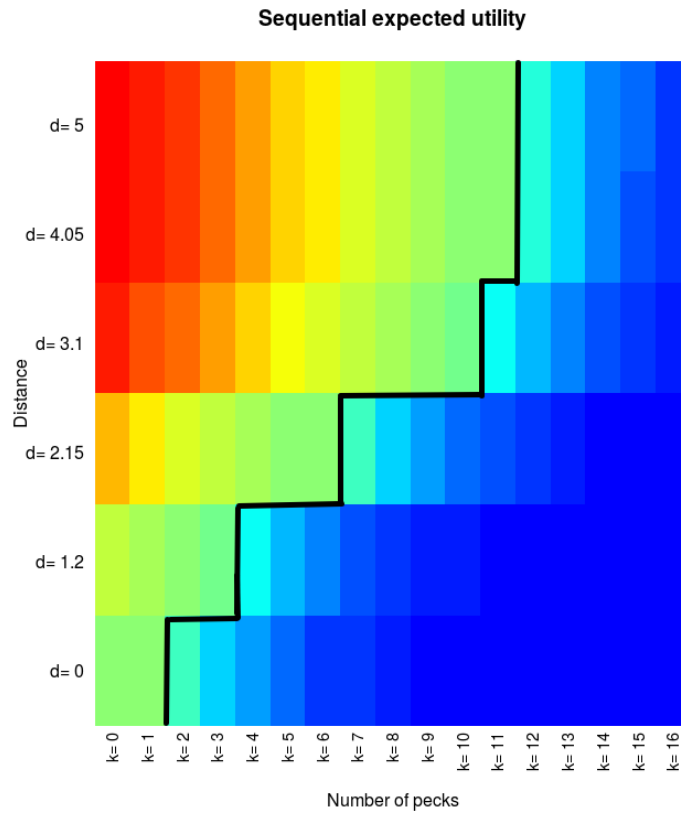


Figure 8: Sequential expected utility for all the choices in the experiment when $r = 50$ and $\sigma^2 = 1$. The utility ranges from 11 in the top-left corner to -1 in the bottom-right corner. The black line separates the choices with a positive sequential expected utility to the left from those with negative to the right. Hence, left of the black line pecking should be done.

positive. If 12 pecks have been performed, it is negative and pecking should not continue. Consequentially, for these parameter values at most 12 pecks should be performed.

4.3 Maximum performance estimation

In the following, we describe how the maximum performance estimator will look in our case. Let us first consider one single trial, and then generalize to several trials. From the definition of performance in (7), we see that we must know how to calculate the chosen expected utility and the optimal expected utility. The former will simply be the sum of the expected utility of all the pecks the bird chose to perform in a trial. Suppose a bird chose to perform k pecks in a trial with $z = z_0$, the chosen expected utility is then given by

$$u(k, \boldsymbol{\theta}) = u_{seq}(0|Z = z_0) + u_{seq}(1|Z = z_0) + \cdots + u_{seq}(k - 1|Z = z_0), \quad (35)$$

where $\boldsymbol{\theta}$ can be any collection of the parameters p_Y , σ^2 , τ^2 , p_N , r and c , depending on what parameters we want to estimate. Note that the terms in (35) can be positive or negative, and that they are in descending order. If the bird did not peck in that trial we let $u(a_c, \boldsymbol{\theta}) = 0$.

As discussed earlier the bird should peck if $u_{seq}(k) > 0$. As the bird pecks, $u_{seq}(k)$ will decrease, when it becomes negative the bird should not peck anymore. Also, if the bird at some point is rewarded it can not peck more. Hence the optimal utility is achieved by pecking either, to the last peck, say peck number i , with $u_{seq}(i) > 0$ or to reward is achieved, whichever happens first. Note that the bird will get a worse chosen utility by pecking more than i times even though it gets rewarded for it. This embodies the fact that the bird does not make good decisions just because it is lucky. Consider the same trial as earlier with $z = z_k$. Suppose $u_{seq}(i|Z = z_k) > 0 > u_{seq}(i + 1|Z = z_k)$, that is peck number i is the last beneficial peck. The optimal expected utility is then

$$u(a_{opt}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \begin{cases} \sum_{j=0}^{j=w} u_{seq}(j|Z = z_k) & \text{if rewarded by peck } w < i, \\ \sum_{j=0}^{j=w} u_{seq}(j|Z = z_k) & \text{otherwise.} \end{cases} \quad (36)$$

The terms in (36) will also be in descending order and they will all be positive as they all are associated with a peck that should be performed. If for some trial there is no pecks with positive expected value, the optimal expected utility is 0. This gain is then achieved by not pecking. Any peck will in such a case result in a negative chosen expected utility.

When considering multiple trials we simply sum up the optimal expected gains for all trials and the chosen expected gain for all trials separately. Suppose we have m trials, the performance become

$$Q(\boldsymbol{\theta}) = \frac{u(a_c, \boldsymbol{\theta})_1 + \cdots + u(a_c, \boldsymbol{\theta})_m}{u(a_{\text{opt}}(\boldsymbol{\theta}), \boldsymbol{\theta})_1 + \cdots + u(a_{\text{opt}}(\boldsymbol{\theta}), \boldsymbol{\theta})_m}, \quad (37)$$

where $u(a_c, \boldsymbol{\theta})_i$ is the chosen gain in trial number i and $u(a_{\text{opt}}(\boldsymbol{\theta}), \boldsymbol{\theta})_i$ is the optimal gain in trial number i . We can then estimate the parameters we are interested in, $\boldsymbol{\theta}$, by maximizing the total performance Q .

4.4 MLE

As discussed in Section 3.2.2, in order to perform MLE we must have a probability distribution describing how probable it is that the bird chooses to peck in the different situations. We established that this probability should be a function of the expected utility alone. Recall that when $u_{\text{seq}}(k|Z = z) > 0$ pecking should be done, while the bird should give up otherwise. The optimal distribution for this experiment is hence to have a probability of zero to peck for negative utility and a probability of one for positive utility, that is a step function. The actual distribution of the bird will naturally deviate from this. The probability of pecking should be low for negative utility, and grow with increasing utility and flat out somewhere below one. Probably the actual distribution, hereinafter referred to as the choice-distribution of the bird, will look like a smoothed step function. This situation is illustrated in Figure 9. The black curve in Figure 9 is called a logistic function, and we choose to model the bird's choice-distribution as this. The logistic function in its general form is very flexible, so we are not assuming much when we make this choice. The probability that the bird will peck when having already performed k pecks is then given by

$$f(a|k) = \begin{cases} \frac{1}{(1+Q \exp\{-B \cdot u_{\text{seq}}(k)\})^{\frac{1}{\nu}}} & \text{if } a = \text{"peck"}, \\ 1 - \frac{1}{(1+Q \exp\{-B \cdot u_{\text{seq}}(k)\})^{\frac{1}{\nu}}} & \text{if } a = \text{"not peck"}, \end{cases} \quad (38)$$

where Q , B and ν are parameters deciding the shape of the logistic curve. Heuristically a rule of thumb is that Q is related to the value of the logistic function at the origin, B is related to the growth rate and ν decides where the most growth is. For simplicity we introduce the notation that

$$A(k) = f(\text{"peck"}|k) = \frac{1}{(1 + Q \exp\{-B \cdot u_{\text{seq}}(k)\})^{\frac{1}{\nu}}}. \quad (39)$$

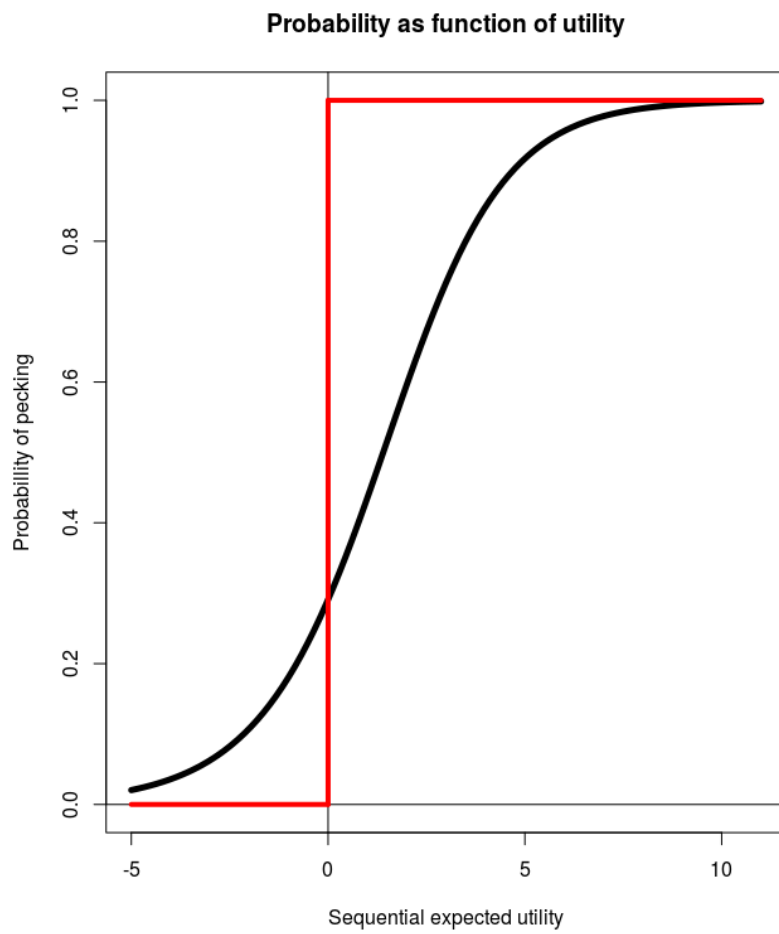


Figure 9: The Optimal distribution in red is a step function. A more realistic distribution is the logistic function in black.

Suppose now we have observed the experiment, and we know how many times the different choices appeared and the decisions of the bird. Specifically let $m_{i,j}$ be the number of times the choice with $d = d_i$ and $k = j$ appeared, $i = 0, \dots, 5$ and $j = 0, \dots, 16$. Furthermore, let $n_{i,j}$ be the number of times the bird choose to peck when $d = d_i$ and $k = j$. Worth noting is that we will always have $m_{i,j} > n_{i,j}$. Also, each peck which do not result in reward will induce another choice with one higher value of k . Since the matching cases never are rewarding we will hence have the identity $n_{0,j} = m_{0,j+1}$. For the matching cases, the bird is sometimes rewarded, so not every peck induces a new decision, instead of equality we will have $n_{i,j} > m_{i,j+1}$. The likelihood function becomes

$$L(r, \sigma^2, Q, B, \nu | m_{i,j}, n_{i,j}) = \prod_{i=0}^{i=5} \prod_{j=0}^{j=16} A^{n_{i,j}} \cdot (1 - A)^{m_{i,j} - n_{i,j}}, \quad (40)$$

where (39) and (26) together shows how the five parameters in the likelihood function is involved. We get the following log-likelihood

$$l(r, \sigma^2, Q, B, \nu | m_{i,j}, n_{i,j}) = \sum_{i=0}^{i=5} \sum_{j=0}^{j=16} n_{i,j} \log(A) + (m_{i,j} - n_{i,j}) \log(1 - A). \quad (41)$$

As this is a very complicated function, we do not try to maximize this by hand. We have experimented with different numeric algorithms and the "Nelder-Mead method" solves the problem quick and precise. This is hence the method we apply. The aim of performing MLE is to arrive at estimates for the five parameters, i.e $\hat{r}_{\text{MLE}}, \hat{\sigma}^2_{\text{MLE}}, \hat{Q}_{\text{MLE}}, \hat{B}_{\text{MLE}}$ and $\hat{\nu}_{\text{MLE}}$. Together these estimates will describe the birds' appreciation of the reward, its evaluation of its own memory and how the probability of pecking depends on the utility, that is, we will have estimated the thoughts of the bird.

However, sometimes two different phenomena cannot be distinguished between. For example, the general relativity theory teaches us that there exists no experiment that can distinguish between gravitation and acceleration. Equivalently in behavioural analysis, it could be the case that two different patterns of thinking may result in the exact same behaviour. As an example, imagine you meet a friend who has cut her hair. The response to tell your friend that you like her new haircut can be motivated by either a genuine opinion that you liked her new haircut or wanting to make a pleasant atmosphere. Hence, when she observes the mentioned response, it is not possible for her to know what your thought process was. Having different patterns of thought is embodied in our experiment as having different parameter values.

Consider two different parameter sets

$$(r_1, \sigma_1^2, Q_1, B_1, \nu_1) \neq (r_2, \sigma_2^2, Q_2, B_2, \nu_2), \quad (42)$$

with the relation $r_1 B_1 = r_2 B_2$. Multiplying this relation with $(\alpha - 1)$, where α is a arbitrary real number, on both sides and adding $(B_1 c - B_1 c) = 0$ to the left side and $(B_2 c - B_2 c) = 0$ to the right side we get

$$B_1(\alpha r_1 - c) - B_1(r_1 - c) = B_2(\alpha r_2 - c) - B_2(r_2 - c). \quad (43)$$

Moving the terms with α to the same side, the terms without α to the other side and let both sides be the exponent of e we get

$$\exp\{B_1(r_1 - c) - B_2(r_2 - c)\} = \exp\{B_1(\alpha r_1 - c) - B_2(\alpha r_2 - c)\} = K. \quad (44)$$

Since the left side does not involve α , the right side must also be invariant of α , and hence equal to a constant K . We now let $K = \frac{Q_1}{Q_2}$, for example by letting $Q_1 = \exp\{B_1(r_1 - c)\}$ and $Q_2 = \exp\{B_2(r_2 - c)\}$. Plugging this in to (44) and moving parameters from the same parameter set to the same side we get

$$Q_1 \exp\{-B_1(\alpha r_1 - c)\} = Q_2 \exp\{-B_2(\alpha r_2 - c)\}, \quad (45)$$

which trivially implies that

$$\frac{1}{(1 + Q_1 \exp\{-B_1(\alpha r_1 - c)\})^{1/\nu_1}} = \frac{1}{(1 + Q_2 \exp\{-B_2(\alpha r_2 - c)\})^{1/\nu_2}}, \quad (46)$$

where $\nu_1 = \nu_2$. Since α was chosen arbitrarily, the latter equation holds for all values of α . From (26) we can see that the utility of pecking can for all the choices be expressed by $\alpha r - c$, by letting α have the appropriate value. We recognize (46) as the probability of pecking from (38). This shows that for two pairs, (r_1, B_1) and (r_2, B_2) , satisfying the relation $r_1 B_1 = r_2 B_2$, and some Q_1 , we can find a Q_2 such that $f(\text{"peck"}|r_1, B_1, Q_1, \nu, \sigma^2) = f(\text{"peck"}|r_2, B_2, Q_2, \nu, \sigma^2)$ for all ν, σ^2 and for all choices, that is all k and d .

We have shown that given some parameter set, there exists another set which gives the same probability distribution for pecking. Since there are infinitely many pairs (r_i, B_i) satisfying $r_i B_i = C$ there are infinitely many parameter sets giving a certain distribution. Furthermore, this obviously implies that the likelihood function evaluated in such equivalent parameter sets will take the same value. Therefore, the likelihood function will not have a maximum point, but instead a curve, described by $r_i B_i = C$ for some C , in the parameter space. This method will consequentially not give a estimate of r nor B ,

only the product rB . To view this in the context of the earlier discussion, we can not distinguish between $(r, B) = (10, 10)$ and $(r, B) = (100, 1)$. Or said differently, suppose two birds follow the probabilistic model in (38), one with a low appreciation of food, but with a rapid growth in probability of pecking with increasing utility will behave identically to another bird with a high appreciation of food, but who needs more utility in order to have the same probability of pecking.

5 Results and discussion

5.1 Maximum performance estimation

In this section, we present the results achieved by using the methods we have discussed. Recall that our aim is to estimate the cognitive parameters and arrive at a probabilistic model for the decision-making.

The maximum performance estimator was implemented according to (37). As discussed in Section 4.2, we use $p_N = 0.3$, $p_Y = 0.8$, $\tau^2 = 5.7075$ and $c = 1$, and estimate r and σ^2 . We use bootstrapping to estimate the standard deviation. Even though 200 bootstrap samples should suffice, we use $B = 1000$ as the algorithm runs fast. The results for Ari from this method is presented in the following table.

Maximum performance estimates for Ari			
	$\theta = r$	$\theta = \sigma^2$	$\theta = Q(r, \sigma^2)$
$\widehat{\theta}_{\text{MPE}}^{\text{Ari}}$	48.326	1.2033	0.83864
$SD_F(\widehat{\theta}_{\text{MPE}}^{\text{Ari}})_{\text{bootstrap}}$	4.3272	0.12134	0.016673

The table shows that, given Ari’s behaviour in the experiment, these parameter values make that behaviour closer to optimal behaviour than any other parameter values. With the idea that the reward is worth $r = 48.326$ and a perceived memory-variance of $\sigma^2 = 1.2033$, Ari’s behaviour achieves 83.864% of the possible expected utility. These estimates suggest that the sequential expected utility Ari is using is closest to the red graph in Figure 7b, where $r = 50$ and $\sigma^2 = 1$. We see that the estimated standard deviation is for both r and σ^2 about 10%, which is quite high. This means that for a slightly different data set, the estimated parameters can be quite different. This suggest that the performance function is fairly flat close to its maximum. The standard deviation for the performance Q is about 5%. Hence,

the estimate of the value of the maximum performance is more precise than the estimated position of the maximum. The corresponding estimates for Susan are given below.

Maximum performance estimates for Susan			
	$\theta = r$	$\theta = \sigma^2$	$\theta = Q(r, \sigma^2)$
$\widehat{\theta}_{\text{MPE}}^{\text{Susan}}$	48.243	1.3984	0.79092
$SD_F(\widehat{\theta}_{\text{MPE}}^{\text{Susan}})_{\text{bootstrap}}$	4.4198	0.15421	0.018733

The estimates indicate that Ari and Susan are about equally eager to be rewarded and that Susan has slightly less belief in her own spatial memory than Ari do. Her behaviour may also be less optimal, as she achieves only 79.092% of the possible expected utility. The standard deviation of Susan's estimates is about the same as those for Ari. The natural question to ask now, is whether or not these estimates are strong enough evidence to claim that there is a difference between the birds. To answer this, we follow the procedure of the hypothesis test explained in Section 3.3. We perform three hypothesis tests, one for each of r , σ^2 and Q . Let the null hypothesis, H_0 , be in each case that there is no difference between the birds. By bootstrapping from the mixed data we get an estimate of the distribution of the estimated difference in a parameter under H_0 . The results are presented in Figure 10. As we can see from Figure 10a, our estimate is in the center of the distribution. It is clear that the estimated difference for r , $\widehat{\delta}_r = 0.083$, is not significant, as expected. There is no reason for believing that this difference does not come from the bootstrap distribution, hence we do not reject H_0 . The estimated differences $\widehat{\delta}_{\sigma^2}$ and $\widehat{\delta}_Q$ are as mentioned leaning in the direction of Ari maybe having a higher belief in own memory and higher performance. However, Figure 10b and 10c shows that the differences are not in the rejection region. We conclude that the MPE-method does not yield sufficient evidence of claiming that there is a difference in the true parameter values between the birds.

A big difference in the birds appear when we look at the correlation between the estimated values for r and σ^2 . The sample correlation for Susan's bootstrap samples is $\text{corr}_{\text{sample}}^{\text{Susan}}(\widehat{r}, \widehat{\sigma}^2) = -0.65$. For Ari we have $\text{corr}_{\text{sample}}^{\text{Ari}}(\widehat{r}, \widehat{\sigma}^2) = 0.02$. This means that for the bootstrap samples of Susan where r is estimated to a high value, σ^2 is more often estimated to a low value. It is not obvious how this should be interpreted. One possible explanation is that Susan sharpens her memory when she is hungry while Ari's memory is on a more constant level. It is both possible that Susan consciously chooses to

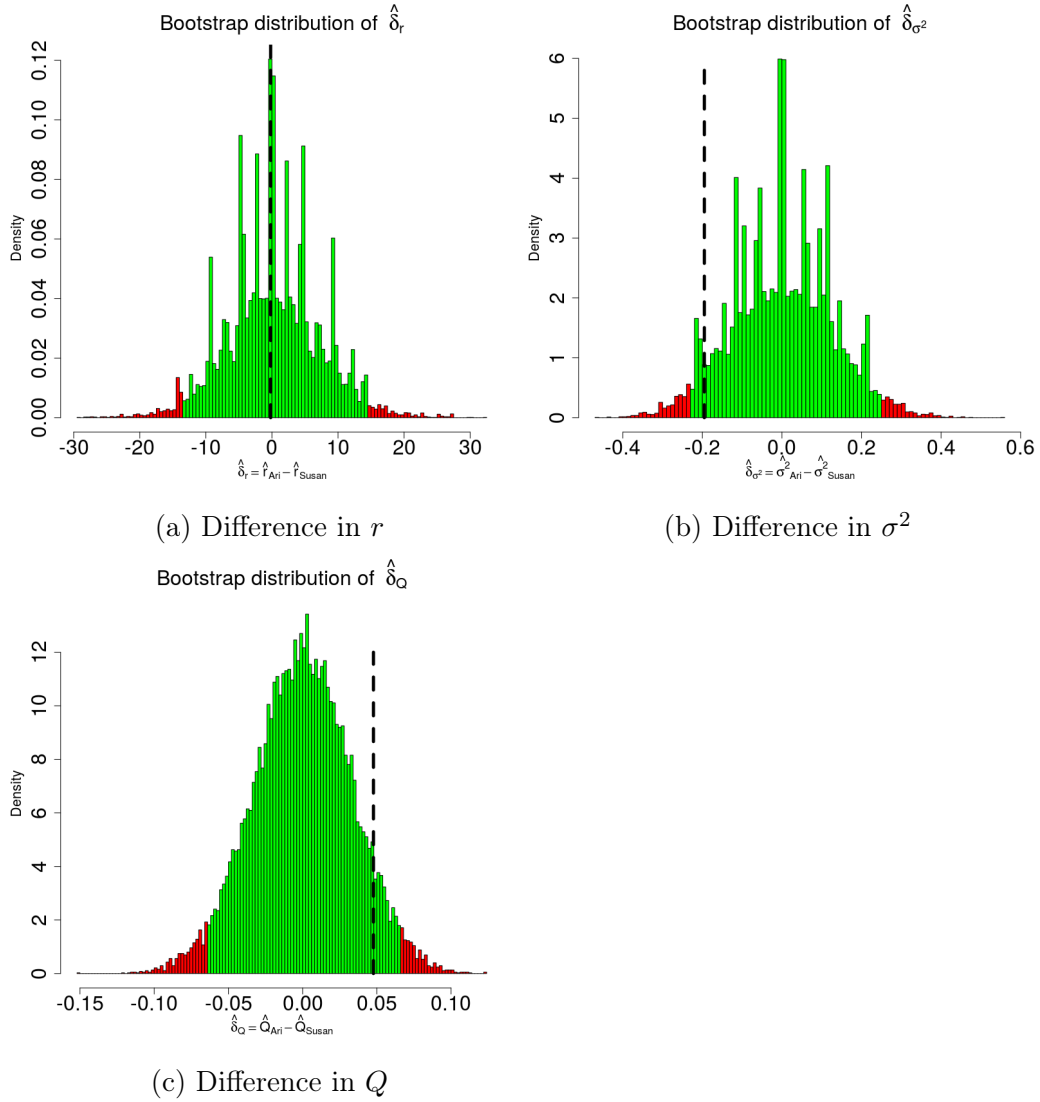


Figure 10: Bootstrap distribution for differences in r , σ^2 and the performance Q . The rejection region is marked with the red area and the acceptance region is the green area. We use the significance level $\alpha = 0.05$. The black dashed line shows the estimated difference between the birds. None of the observations are in the rejection region. Note that we always use $\delta_\theta = \theta_{Ari} - \theta_{Susan}$.

work harder since she wants food more, or that this is a result of a subconscious survival mechanism. It would be interesting to conduct an experiment to further investigate whether Ari and Susan responds differently to being hungry.

5.2 Maximum likelihood estimation

We now present the results from the maximum likelihood estimation. Recall that in addition to using another estimation method, we are also using another model which includes a choice-distribution that models a probabilistic behaviour. As we discussed in Section 4.4, we can only estimate the product rB , not r and B separately. Choosing some value for B will influence the estimated value of r and Q , but nothing else. The fitted model describing how the birds behave will be the same for all choices of B . For simplicity let $B = 1$ for both birds. Also for these estimates, we use bootstrapping to estimate the uncertainty. The number of bootstrap samples was as above $B = 1000$. The resulting estimates for both birds are given in the tables below.

Maximum likelihood estimates for Ari				
	$\theta = Q$	$\theta = \nu$	$\theta = rB$	$\theta = \sigma^2$
$\widehat{\theta}_{MLE}^{Ari}$	3.1606	5.3447	19.304	0.64627
$SD_F(\widehat{\theta}_{MLE}^{Ari})_{bootstrap}$	15.515	2.3809	3.8897	0.091367

Maximum likelihood estimates for Susan				
	$\theta = Q$	$\theta = \nu$	$\theta = rB$	$\theta = \sigma^2$
$\widehat{\theta}_{MLE}^{Susan}$	0.30772	1.6578	11.360	0.61550
$SD_F(\widehat{\theta}_{MLE}^{Susan})_{bootstrap}$	1.1410	1.9516	2.6097	0.18028

We see that for both birds, the MLE-estimates of σ^2 is much smaller than the corresponding MPE-estimates. However, when we change our model, the parameter interacts. This means that we can not expect the value of σ^2 in this model to be the same as before introducing the three parameters linked to the choice-distribution. The tables show that the three leftmost estimates differ a lot between the birds. However, due to the high uncertainty asso-

ciated with the estimates, the difference might be statistically insignificant. Note that the estimate of Ari's Q -value has a substantial estimated standard deviation of 15.515. This might seem unreasonable. However, recall that the logistic function compresses all the real numbers in to the interval $(0, 1)$. The value of Q is very sensitive to changes in the choice-distribution. Also, the effect the three parameters have on the shape of the function do to some extent overlap. This means that two functions which seems graphically very similar can have fairly different parameter values. Hence when we perform bootstrap sampling and change our data set a bit, we might see big changes in the estimated parameter values. From the tables, we can see that $\widehat{\delta}_{rB}$ is the only difference that is bigger than the sum of the corresponding estimated standard deviation for each bird. This might therefore be the best candidate of being a statistical significant difference. As before, we perform hypothesis tests to decide whether the estimated differences in parameter values are significant. As before, let the null hypothesis be that there are no difference, and that the true value of the parameters are the same for the birds. The bootstrap distributions of the differences under H_0 are plotted in Figure 11. From the figure we can see that even though there are big estimated differences, none of them are statistically significant. The observed differences could very well appear even if the birds were identical. Clearly the estimated parameter values are very sensitive to changes in the data set. When we sample with replacement from the original data set we might get very different estimated parameter values. Note that the difference in rB is the closest one to being in the rejection region. This is coherent with our expectation that it was the best candidate of being a statistical significant difference.

The fitted choice distributions associate probability of pecking to expected utility. This can be compared to the observed amount of pecking used to fit the model. All 6 · 17 choices in the experiment has an associated utility given by the fitted model. For each choice we also have observed in what proportion the birds chose to peck. If the model is realistic that proportion should not be too different from the choice distribution evaluated in the expected utility of that choice, at least for the choices which appeared many times. For a choice with a probability of 0.6 of pecking according to the model, we might very well find the observed proportion to be 0 if the choice appeared only a few times in the experiment. The estimated choice distribution together with the observed proportions are given in Figure 12 for Ari and in Figure 13 for Susan. As we can see from the figure for Ari, the true proportions are not lying too close to the fitted model. This alone need not be alarming. When sampling we must expect that the proportions will deviate from

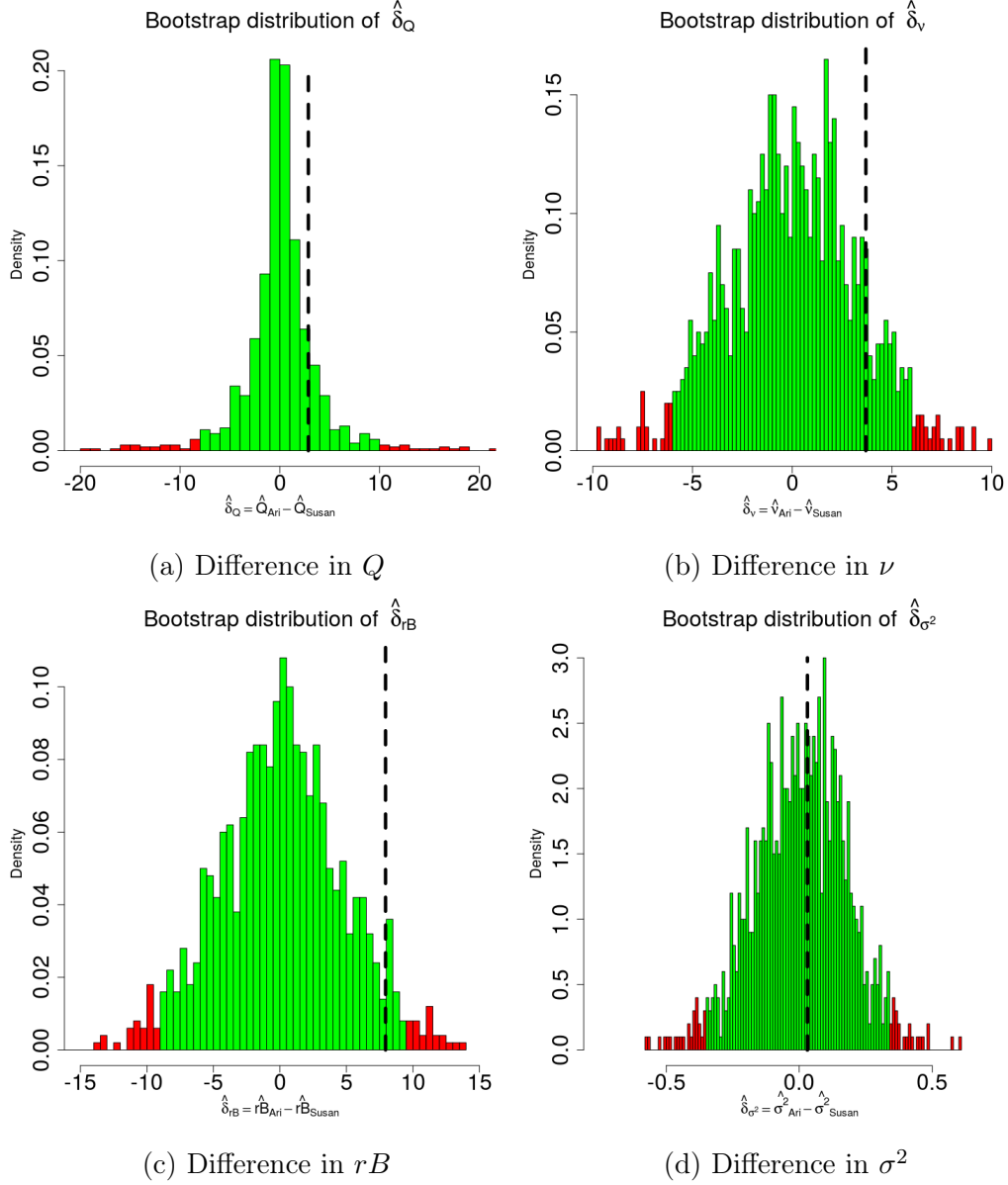


Figure 11: Bootstrap distribution for differences in Q , ν , rB and σ^2 . The rejection region is marked with the read area and the acceptance region is the green area. We use the significance level $\alpha = 0.05$. The black dashed line is the estimated parameter difference between the birds. None of the observations are in the rejection region. Note that we always use $\delta_\theta = \theta_{Ari} - \theta_{Susan}$.

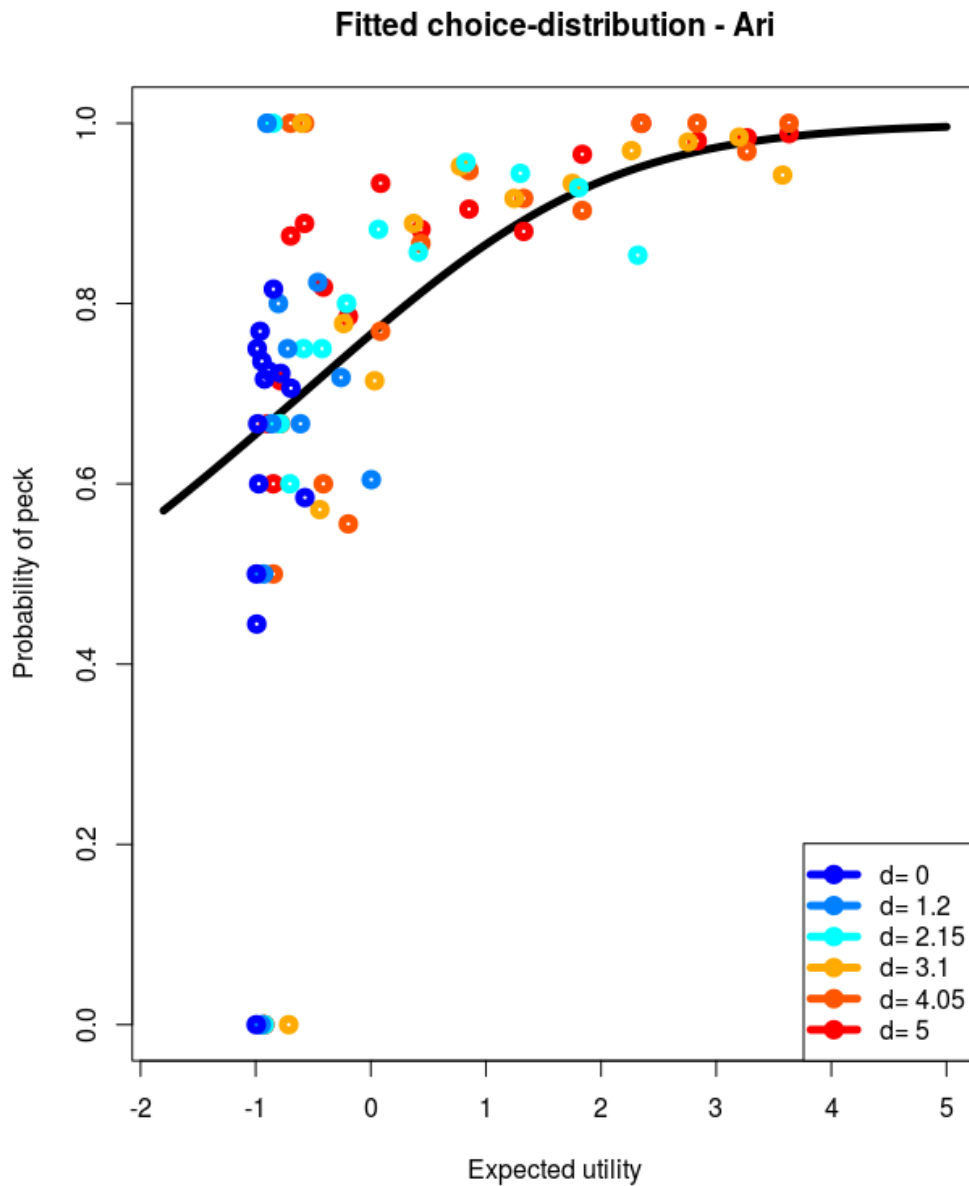


Figure 12: The black curve shows the estimated relation between expected utility and probability of peck. Each dot corresponds to a choice in the experiment, where the expected utility is calculated using the MLE-estimates and the second coordinate is the observed proportion of pecking in the experiment.

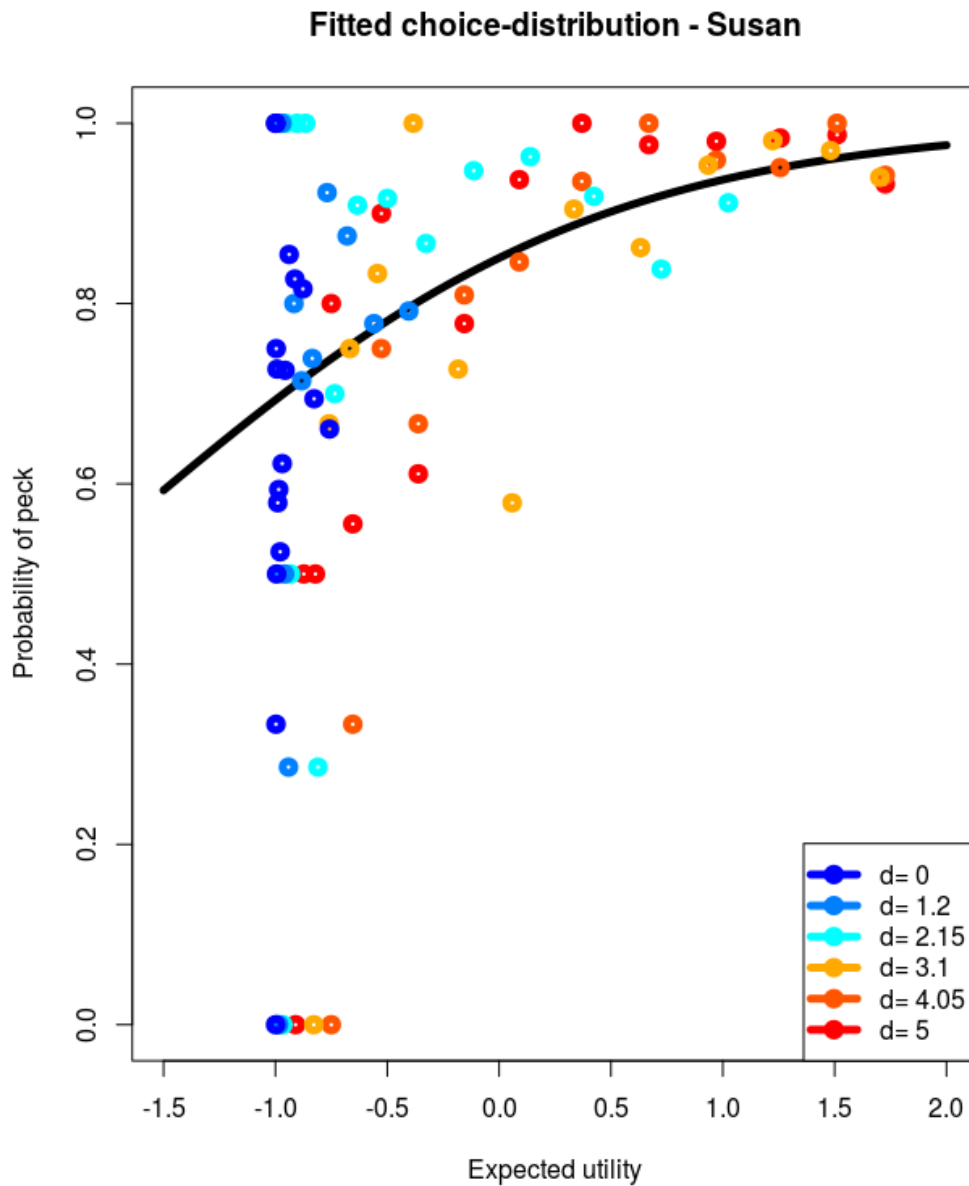


Figure 13: The black curve shows the estimated relation between expected utility and probability of peck. Each dot corresponds to a choice in the experiment, where the expected utility is calculated using the MLE-estimates and the second coordinate is the observed proportion of pecking in the experiment.

the probability. However, if the model was realistic the deviation from the black curve should seem random and noisier. It seems as if it is a systematic difference. The points which are below the curve are over-represented for low utilities. It might seem as if the model does not fit very well. On the other hand, this evaluation is purely qualitative and does not give reason for rejecting the model yet. For Susan's figure, the deviation from the black curve seems more evenly spread. This suggests that the model is fitting her better than Ari. Recall that some of the dots in the figures carry much more weight than others. The rightmost dark blue dot represents the choice with $d = 0$ and $k = 0$ which appeared 419 times for Ari and 510 for Susan. Some choices only appeared once. Therefore, the black curve will not necessarily appear visually to be the best fitting curve.

Recall from the discussion in Section 3.2.2 that the probability of pecking should depend only on the associated expected utility. If we had in Figure 12 and 13 that the different colours were in layers, then this would indicate that our utility function is unrealistic. If choices with a certain distance systematically were below choices with another distance, then our utility function does not account for the value of d in a realistic manner. Hence, considering whether such a pattern is apparent in the figures is a qualitative test for our utility function. Looking at the figures, we see no such pattern. No colour is systematically above any other colour. This gives an indication that the utility is not unrealistic. We do see colour grouping in that the blue dots are only on the left side, and the right side is dominated by orange and red dots. This is, however, to be expected. There are no choices with high expected utility for $d = 0$, as these trials never are rewarding.

To more easily compare the choice-distribution of the two birds they are illustrated together in Figure 14. The figure shows that the red curve is mostly above the blue. The curves do cross at $u_{seq} = 1.5$, but recall that no choices have lower expected utility than $-c = -1$. Also, as Figure 12 and 13 show, all choices have lower expected utility than 2. Hence for all expected utilities associated with a choice, Susan has a higher pecking probability. This might lead one to believe that Susan is in general pecking more than Ari does. When asserting a probability of pecking to a choice, our model first associates a utility to that choice and then a probability to that utility. It is, in other words, a two-step process. If a bird pecks a lot, this can be described by both of these steps. A bird that has a very high utility for all choices can have a slow increasing choice-distribution and still peck a lot. A good example of this is that only the product rB is relevant for the behaviour. A high value for r contributes to a high utility for choices, while a high value

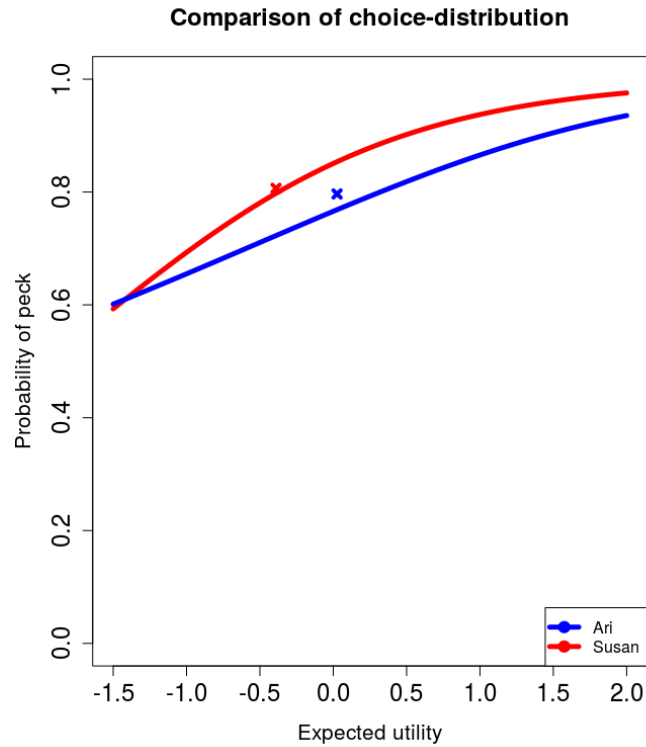


Figure 14: Probability of peck as a function of expected utility for the fitted model. This is also referred to as the choice-distribution. The first coordinate of the crosses indicates average expected utility where each of the $6 \cdot 17$ possible choices is weighted equally, not accounting for the fact that some appeared more in the experiment. The second coordinate of the crosses is the number of pecks performed divided by the number of opportunities to peck. The red curve and cross represent Susan and the blue curve and cross represent Ari.

of B does not. That value will instead contribute to a high peck probability as a function of utility. Recall that when using $B = 1$ we have estimated $\hat{r}_{\text{Ari}} = 19.304$ and $\hat{r}_{\text{Susan}} = 11.360$. Because of this, Ari has typically a higher utility for a choice than Susan has. But for a certain utility, Susan has a higher probability. These two effects are opposing and evening each other out. As the crosses in Figure 14 shows, Susan is pecking marginally more on average than Ari does. The difference is however much smaller than what the curves might make it seem like, due to the discussed effect.

5.3 Simulation

From the probabilistic model for the birds' behaviour, we can generate new data by simulating experiments. Simulated data is very interesting since we know for sure that they are described by the model. Comparing generated behaviour to the observed behaviour is a powerful method for answering whether it is realistic that the observed behaviour comes from this model.

Simulation also allows us to test our estimators. When simulating data we know the true value of the parameters since we choose them ourselves. Therefore, estimation based on generated data will show how good the estimators perform. Both the variance and whether or not the estimator is unbiased can be investigated by this technique. To keep the analysis brief we only do this for Ari. We generate 1000 samples from Ari's fitted model, each of sample size 846 trials. For each of the samples we estimate the four parameters Q , ν , rB and σ^2 . The estimates will then make up an approximated distribution for the estimators. The results are shown in Figure 15. The distribution of Q has a big weight close to zero and a long tail. Due to the long tail, we see that the estimator becomes pretty unbiased with an average estimation of about 20. However, the median of the distribution is 3.44 which is fairly close to the used value 3.16. The estimator is definitely unbiased in expectation, but this is only due to some very few very extreme observations. There are for example 3 observations over 1000. Looking at the distribution in Figure 15a we see that the estimate has a high probability of being close to the true value. The reason we test our estimators is to see whether our estimates are realistic, that is, close to the true values. We prefer an estimator that is biased but almost always close to the true value, and not an estimator that is unbiased but almost always far from correct. Hence we do not use bias correction. The spread of this distribution shows, as we have already seen, that the variance in the estimation of Q is very big. All MLE-estimators are consistent. This means that they converge to the true value. Using a sample size of 2000 instead of 846 the average estimate of Q

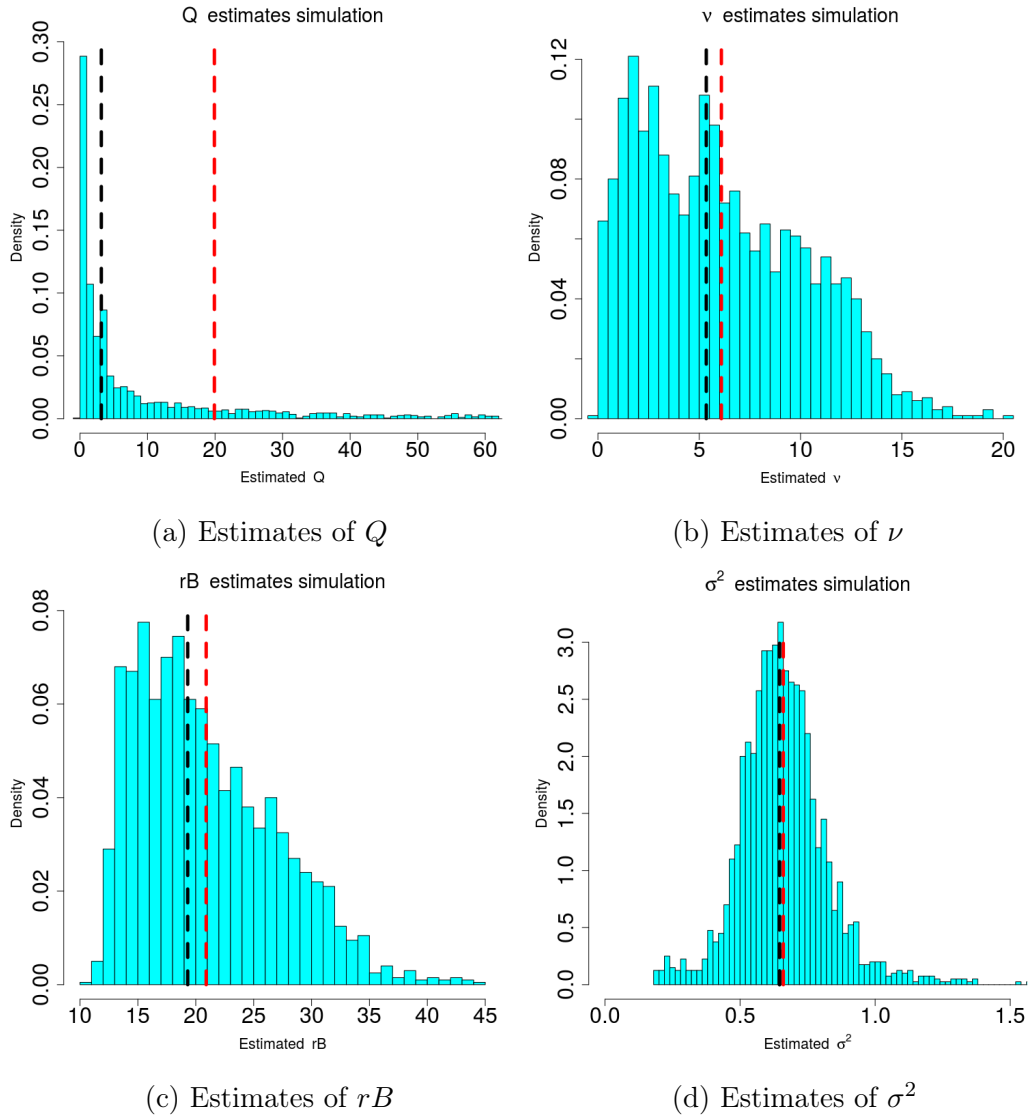


Figure 15: Estimates of the four parameters based on 1000 generated data sets of 846 trials from Ari's fitted model. The black dashed line is the used parameter value in the simulation. The average estimated values is the red dashed line.

Giving Up Without Pecking - Simulation

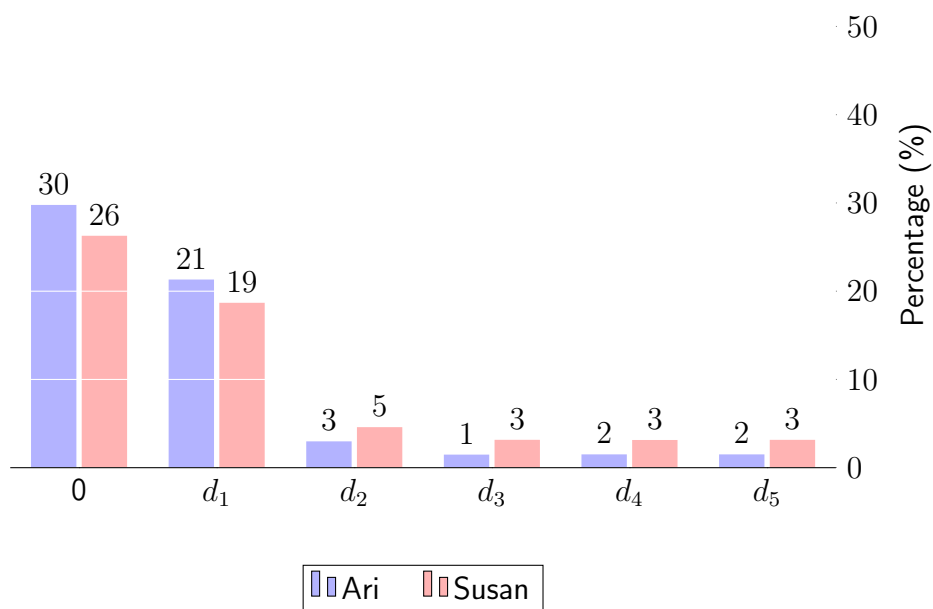


Figure 16: Percentage of simulated trials the birds gave up without pecking, based on 100,000,000 generated trials. The figure shows a blue bar for Ari and a pink bar for Susan for six different values of d . The matching cases are represented in the leftmost bar-pair where $d = 0$. The non-matching cases make up the five other bar-pairs.

becomes 7.44. A further increase to 10,000 trials gives 3.62. This reflects the consistency of MLE. The estimators of ν , rB and σ^2 all average close to the used parameter value, as shown in Figure 15.

So far we have used simulated data to test our estimators. We will in the following compare behaviour simulated from the model to the observed behaviour. If the simulation deviates a lot from the observed data, then this indicates that the model assumptions are unrealistic. We first simulate 100,000,000 trials for each bird and illustrate the behaviour in the same way as we did with the observed data. First we look at the giving-up percentage, shown in Figure 16. The general trend in the figure compared to Figure 2 is that the birds give up less in the simulation than in the experiment. The exception from this trend is Ari's extreme low giving-up percentage at the two biggest d values, which is even lower than the simulated ones. We see that Susan's simulated behaviour is to give up less for all distance levels and at d_5

Average Investment - Simulation

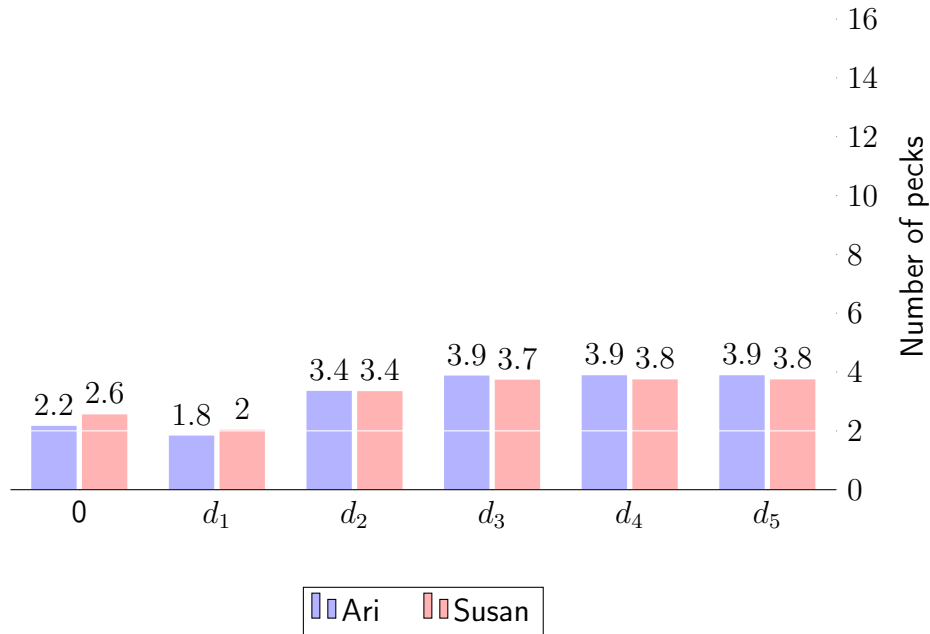


Figure 17: The average number of pecks for each distance level is illustrated, including the trials where the bird gave up, i.e. pecked zero times. We use 100,000,000 simulated trials. The blue bars represent Ari’s simulated investments and the pink bars represent Susan’s simulated investments.

she is giving up over twice as much in the experiment than in the simulated data. Perhaps the biggest difference between reality and the simulation is Ari’s behaviour at $d = 1$. In the experiment, he gave up 40% compared to the simulated 21%. This difference seems pretty big and may lead us to believe that the model is unrealistic. This is simply a qualitative observation and is not sufficient evidence for rejection of the model. We will later perform hypothesis tests to see if the deviation from the model is significant.

We now consider the average investments at different distance levels. This is illustrated in Figure 17. The figure matches the observed investments in Figure 3 quite well in several aspects. As the simulated data shows, the model reproduces the fact that the average investment for both birds is lower at d_1 than d_0 . As discussed earlier, this is due to the birds being rewarded at d_1 , and thus hindered from pecking as much as they are willing to. We also see that the fitted models have caught up the fact that Ari is pecking less than Susan for low values of d , but that it evens out for bigger values of d . Overall

Average Investment in Non-Rewarding Trials - Simulation

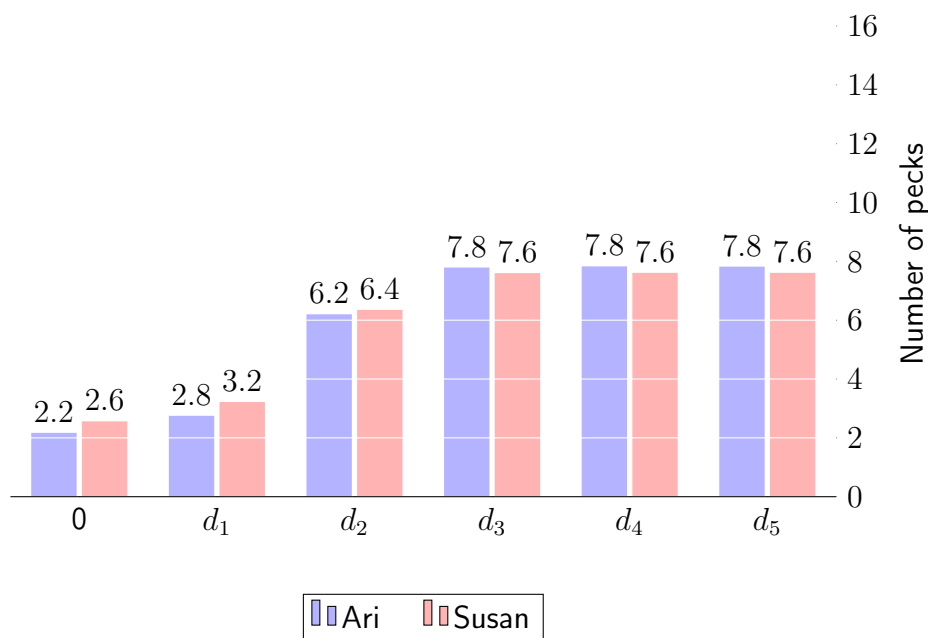


Figure 18: The average number of pecks for each distance level when only considering non-rewarding trials is illustrated. The data comes from a simulation of 100,000,000 trials, resulting in about 20,000,000 non-rewarding trials. The blue bars represent Ari’s simulated investments and the pink bars represent Susan’s simulated investments.

this figure matches the observed data well.

To see the investments when the birds are not being hindered by being rewarded, we have in Figure 18 plotted the average investment in non-rewarding trials. As expected the average investments are increased when only considering non-rewarding trials. Also, note that the average investment at d_1 becomes bigger than the average investment at 0. This effect is coherent with the observed data in Figure 4. Both birds peck more at d_5 in reality than in the simulations, especially Ari. Ari was also giving up a lot more for low values of d than the simulations suggest. It might seem that Ari is not as willing to start pecking as often as the model suggest, but when he first starts pecking, he is willing to peck more than the model. This suggests that the model assumption that each peck costs the same is unrealistic. Perhaps a better assumption for Ari would be that the first peck in a trial has a bigger cost.

Another plot that is very descriptive of the behaviour is the distribution of the number of pecks performed in a trial. The observed distribution from the experiment is plotted together with the distribution from the simulation in Figure 19. The figure shows that the simulated data matches the observed data pretty good. For both birds, we have that both the observed data and the simulated data very seldom show more than 15 pecks. All four densities have a peak of about 0.25 which is at either 0 pecks or 1 peck. In fact, the most obvious difference between the simulated data and the observed data is the relation between the 0 and 1. The model suggests that the proportion of trials where 1 peck is performed should be at least 5 percentage points bigger than the proportion with 0 pecks. This is not observed for Ari nor Susan. Ari even has the opposite relation between 0 peck and 1 peck. The figure supports our suspicion that Ari, in reality, has a bigger cost for the first peck than the model does.

Hitherto, we have only compared the simulated behaviour and the observed behaviour qualitatively. We have pointed at some trends that might give reason to believe that the model is unrealistic. For example, we noted us that the giving-up percentage for Ari at d_1 and for Susan and d_5 was quite different from that of the simulated data. We will now do some quantitative analysis such that we can consider whether it is evidence strong enough to reject the proposed model. As before, we perform hypothesis tests. Let the null hypothesis, H_0 , be that the observed data is generated from the model. Let us consider a couple of test statistics and estimate their distributions by simulating data. If a test statistic of the observed data is unreasonably far out in the tail of the estimated distribution, we conclude that the observed data does not have the same distribution as the model does. That is, H_0 is rejected. A very natural test statistic to use is the average number of pecks in a trial for a data set of size 846 for Ari and 1016 for Susan. This statistic summarizes and weights all trials equally, its estimated distribution, using 1,000,000 datasets, is plotted in Figure 20. The figure shows for both birds that the observed average number of peck is near the middle of the density. Hence we do not reject H_0 .

Let us now use some test statistics that we suspect will result in rejection. Recall that Ari seemed to give up unreasonably much in trials with $d = d_1$. Hence, we perform a hypothesis test where the test statistic is the giving-up percentage at d_1 . Again, the null hypothesis is that the observed data is generated from the model. For this test we also use 1,000,000 data sets of size 846 for Ari and 1016 for Susan. The result is illustrated in Figure 21. As

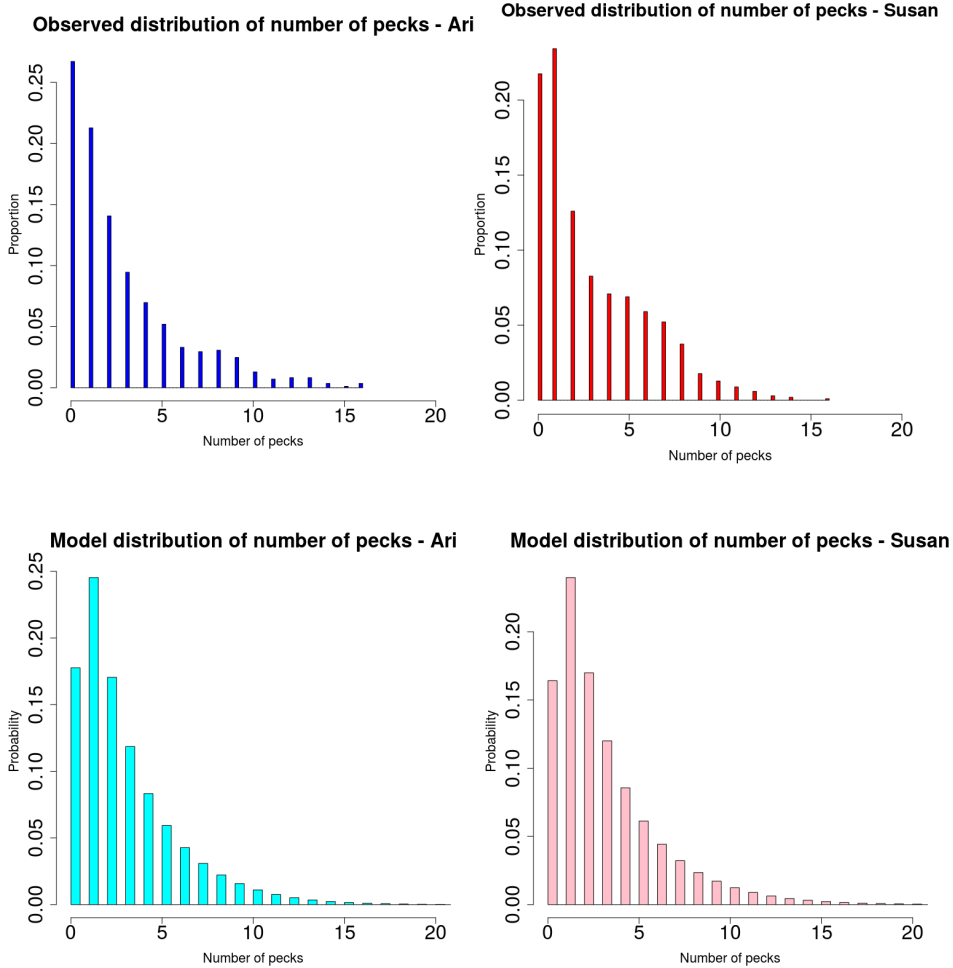


Figure 19: Distribution of number of pecks performed in trials, based on 100,000,000 trials.

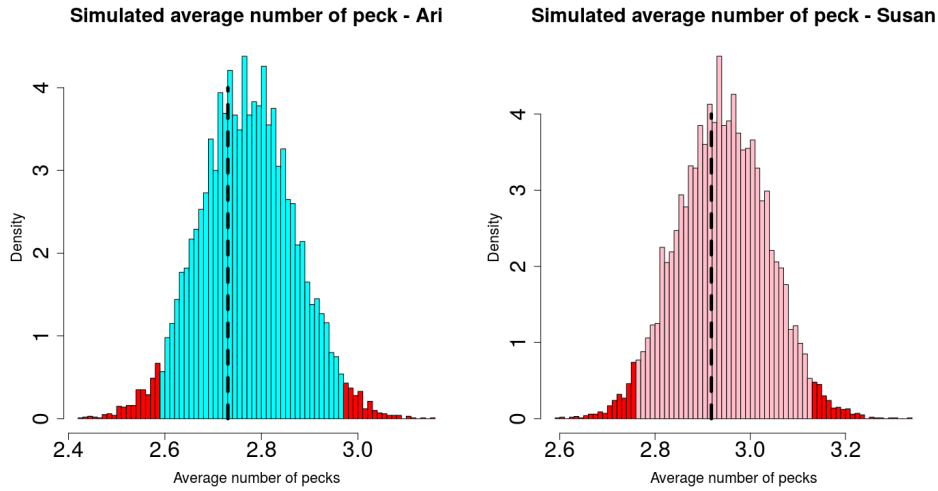


Figure 20: Distribution of the average number of pecks in simulated data sets of size 846 for Ari and 1016 for Susan. For both birds we have used 1,000,000 data set, each used to compute one average. The red area represents the rejection region of the hypothesis test. The black dashed line is the observed average.

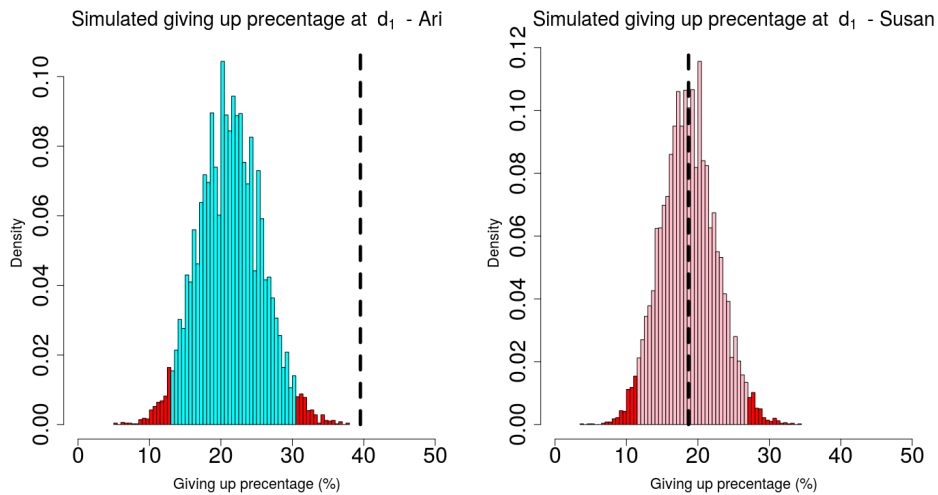


Figure 21: Distribution of the giving-up percentage at d_1 , based on 1,000,000 simulated data sets of the birds' respective data set size. The red area is the rejection region and the black dashed line is the observed giving-up percentage.

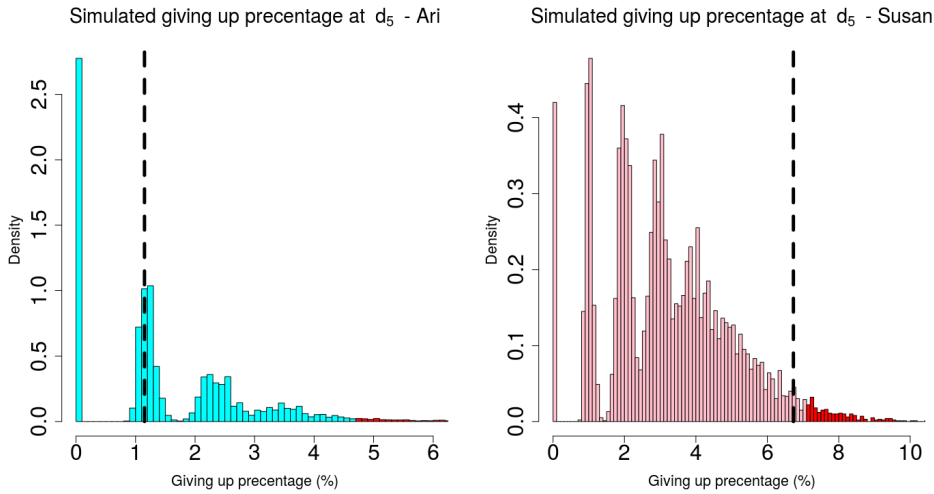


Figure 22: Distribution of the giving-up percentage at d_5 , based on 1,000,000 simulated data sets of the birds' respective data set size. The red area is the rejection region and the black dashed line is the observed giving-up percentage.

the figure shows, the observed giving-up percentage at d_1 for Ari is far out in the tail of the distribution and definitely in the rejection region. This aspect of Ari's behaviour can not be explained by the model. Hence we reject H_0 for Ari. In contrast, Susan's giving-up percentage at d_1 fits the model very well. Recall that we criticized the model for Susan regarding her high giving-up percentage at d_5 . We now perform the same hypothesis test at the distance level $d = d_5$. The result is illustrated in Figure 22. From the figure we see that neither of the birds' observed giving-up percentage at d_5 is extreme enough to reject H_0 . Note however that Susan's observation is very close to the rejection region, which is coherent with our suspicion. The distributions might appear to look strange. However, recall that both birds give up very seldom at d_5 . Therefore we can see local maximums corresponding to giving up one time, two times and three times. The reason for the spreading around these maximums is that the number of d_5 trials in a data set is stochastic. Note that neither of the bird has a rejection region on the left side. The lowest possible observation is 0%, which is not extreme enough to be in the rejection region.

6 Closing remarks

In this thesis, we have considered an experiment on metacognition in jay-birds. We established a statistical model for the experiment and investigated how the optimal investment pattern depends on the parameters of interest, the reward r and the perceived spatial memory variance σ^2 . We introduced and applied maximum performance estimation based on the observed data to estimate r and σ^2 . The results showed that the birds are equally eager to be rewarded with food. The estimated value for σ^2 was slightly higher for Susan than for Ari. However, by using bootstrapping to perform a hypothesis test, we saw that this difference is not statistically significant at a $\alpha = 0.05$ -level. The performance of Ari was significantly higher than Susan's. In our model, this should be interpreted as Ari making better decisions than Susan does. The bootstrap resampling also revealed that the estimated values for r and σ^2 for Susan are correlating. The correlation factor of -0.65 could suggest that Susan has more belief in her own memory when hungry. Ari has no such correlation, and it would be interesting to conduct an experiment to further investigate whether Susan and Ari really do respond differently to being hungry.

Another approach was used where we established a full probabilistic model for how the birds make decisions. The same utility function was used, but for a certain utility, we associated with it a probability of choosing to peck, by defining a choice-distribution. As this model gives the probability that some decision is made, it allowed us to use maximum likelihood estimation. The use of a choice-distribution adds three more parameters, Q , B and ν , describing the shape of the choice distribution. We argued that r and B could not be estimated, only their product. When using MLE we found big differences between the birds in the estimated parameter values. However, by performing a bootstrap hypothesis test we saw that these differences were not significant. From the fitted model, we simulated new data for comparison with the observed data. In a lot of aspects, the model describes the behaviour of the birds well. Based on the generated data we performed three hypothesis tests for each bird. The observed behaviour of Susan did in none of these tests fall in the rejection region. Hence, we have not rejected that she is behaving like the model suggests.

The fitted model for Ari did not pass all hypothesis tests. His giving-up percentage of 40% at d_1 showed to be significantly different from the simulated 21%. Hence we reject the hypothesis that Ari's behaviour is generated from this model. This means that some model assumptions are not fitting

Ari's behaviour. There are several indicators pointing toward the cost Ari experiences by performing the first peck in a trial is bigger than in the model. First, we saw by comparing Figure 16 and 2 that in the simulations, Ari is not giving up as much as in reality. This was later supported by considering the distribution of the number of pecks, as shown in Figure 19. Lastly, the aspect of his behaviour that makes us reject Ari's model, his giving-up percentage at d_1 , would also in part be explained by a higher cost for the first peck.

It seems to be reasons to believe that the birds respond differently to being hungry. It would, therefore, be interesting in further work to examine whether this could be the case, for example by conducting an experiment suited for that. Furthermore, since our model is not fully describing the behaviour of the birds, we suggest the usage of a more sophisticated utility function. By not assuming each peck having the same cost, Ari's behaviour is likely to be better explained.

References

Macody Lund S. (2018). Analysis of an Investment Problem for an Experiment with Jaybirds. Project report, TMA4500 Specialization project, Dept. of Mathematical Sciences, Norwegian University of Science and Technology.

Herrnstein R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior* Volume 4; 267–272.

Borrero J.C, Crisolo SS, Tu Q., Rieland W.A., Ross N.A., Francisco M.T., Yamamoto K.Y. (2007). An application of the matching law to social dynamics. *Journal of Applied Behavior Analysis* Volume 40; 589-601.

Pfuhl G., Biegler R. (2012). Do jackdaws have a memory for order? *Journal of Articles in Support of the Null Hypothesis* Volume 9 No 1. 25 pages.

Fardzadeh H.E, Heydari A, Biegler R, Pfuhl G, Teymoori A. (2012) Sociological Indicators of Ethical Attitude toward Ethnic Humor: An Empirical Study on Persian Jokes. [*Asian Social Science*] Volume 8 No 13; 9-20.

Walpole R.E., Myers R.H., Myers S.L., Ye K. (2012) Probability & Statistics for Engineers & Scientists, 9th edition; 307-312.

Givens G.H. and Hoeting J.A. (2013). Computational Statistics, 2nd edition; 287-319.

Appendices

A Calculation of $E[YI(X \neq Z)I(N = k + 1)r - c|Z = z, Y = 0 \cup X = Z \cup N > k]$

We now present the calculation of $E[YI(X \neq Z)I(N = k + 1)r - c|Z = z, Y = 0 \cup X = Z \cup N > k]$ as mentioned in Section 4.2. The expression $YI(X \neq Z)I(N = k + 1)$ is equal to 1 if and only if all three factors are equal to 1, otherwise the expression is zero. Hence, by using the fact that $E[I(A)] = 1 \cdot P(A) + 0 \cdot (1 - P(A)) = P(A)$, we get

$$\begin{aligned} & E[YI(X \neq Z)I(N = k + 1)r - c|Z = z, Y = 0 \cup X = Z \cup N > k] \\ &= P(Y = 1, X \neq Z, N = k + 1|Z = z, Y = 0 \cup X = Z \cup N > k) \cdot r - c. \end{aligned} \quad (47)$$

By using the definition of conditional probability we can write

$$\begin{aligned} & P(Y = 1, X \neq Z, N = k + 1|Z = z, Y = 0 \cup X = Z \cup N > k) \\ &= \frac{P(Y = 1, X \neq Z, N = k + 1, Y = 0 \cup X = Z \cup N > k|Z = z)}{P(Y = 0 \cup X = Z \cup N > k|Z = z)}. \end{aligned} \quad (48)$$

The triple union in the numerator can be cancelled since it is intersected with the event $N = k + 1$ which implies $N > k$. The remaining events in the intersection are independent of each other and only the event $X \neq Z$ depends on the condition $Z = z$. Also we use that $P(A \cup B) = P(A) + P(B \setminus A)$, where $B \setminus A$ is the set difference, two times in the triple union in the denominator and get

$$\begin{aligned} & \frac{P(Y = 1, X \neq Z, N = k + 1, Y = 0 \cup X = Z \cup N > k|Z = z)}{P(Y = 0 \cup X = Z \cup N > k|Z = z)} \\ &= \frac{p_Y P(X \neq Z|Z = z)(1 - p_N)^k p_N}{(1 - p_Y) + p_Y P(X \neq Z|Z = z) + p_Y P(X \neq Z|Z = z)(1 - p_N)^k}. \end{aligned} \quad (49)$$

We see from this expression that it only remains to calculate $P(X \neq Z|Z = z)$, that is the probability from a bird's perspective that the initial dot's position is not equal to the final dot, which it knows the position of. This is in other words the perceived probability of a non-matching trial. We first write this in terms of a integral of the density

$$P(X \neq Z|Z = z) = \int_{x \neq z} f(x|z)dx = 1 - \int_{x=z} f(x|z)dx. \quad (50)$$

Furthermore, Bayes theorem yields

$$f(x|z) = \frac{f(x)f(z|x)}{f(z)} = \frac{f(x)f(z|x)}{\int_{x \in \mathcal{D}} f(x)f(z|x)dx}, \quad (51)$$

in the latter equality we use the law of total probability. The denominator in (51) is a constant. hence we only integrate the numerator

$$\int_{x=z} f(x)f(z|x)dx = \int_{x=z} n(x; \mu, \sigma^2) \left(\frac{1}{2}\delta_x(z) + \frac{1}{2}n(z; x, \tau^2) \right) dx. \quad (52)$$

An infinitesimal size multiplied with some finite size is an infinitesimal size, and integrated over a single point it evaluates to zero. For the term involving the dirac delta function we use that $\int_{\mathcal{R}} \delta_{x_0}(x)(g(x)dx = g(x_0)$ for all functions g on \mathcal{R} , we get

$$\int_{x=z} n(x; \mu, \sigma^2) \left(\frac{1}{2}\delta_x(z) + \frac{1}{2}n(z; x, \tau^2) \right) dx = \frac{1}{2}n(z; \mu, \sigma^2). \quad (53)$$

Next we consider the denominator in (51), substituting in the densities we get

$$\int_{x \in \mathcal{D}} f(x)f(z|x)dx = \int_{x \in \mathcal{D}} n(x; \mu, \sigma^2) \left(\frac{1}{2}\delta_x(z) + \frac{1}{2}n(z; x, \tau^2) \right) dx. \quad (54)$$

The dirac delta function is equal to zero for all values other than $x = z$, hence the first term integrates to the same as (53), we can thus write

$$\begin{aligned} \int_{x \in \mathcal{D}} n(x; \mu, \sigma^2) \left(\frac{1}{2}\delta_x(z) + \frac{1}{2}n(z; x, \tau^2) \right) dx = \\ \frac{1}{2}n(z; \mu, \sigma^2) + \frac{1}{2} \int_{x \in \mathcal{D}} n(x; \mu, \sigma^2)n(z; x, \tau^2)dx. \end{aligned} \quad (55)$$

Due to the law of total probability the latter integral is the density of z given that $X \neq Z$. Sampling $X \sim N(\mu, \sigma^2)$ and from that realization x_0

sampling $Z \sim N(x_0, \tau^2)$ is equivalent with sampling $Z = X_1 + X_2$, where $X_1 \sim N(\mu, \sigma^2)$ and $X_2 \sim N(\mu, \tau^2)$. Hence the density of z in the non-matching cases is the density of $X_1 + X_2$. The random variables X_1 and X_2 are independent, thus the variance of Z is the sum of the variance of X_1 and X_2 . That is

$$\int_{x \in \mathcal{D}} n(x; \mu, \sigma^2) n(z; x, \tau^2) dx = n(z, \mu, \sigma^2 + \tau^2). \quad (56)$$

This means that the denominator in (51) is

$$\int_{x \in \mathcal{D}} f(x) f(z|x) dx = \frac{1}{2} n(z; \mu, \sigma^2) + \frac{1}{2} n(z; \mu, \sigma^2 + \tau^2). \quad (57)$$

By plugging this and (53) into (50) we can write

$$\begin{aligned} P(X \neq Z | Z = z) &= 1 - \frac{\frac{1}{2} n(z; \mu, \sigma^2)}{\frac{1}{2} n(z; \mu, \sigma^2) + \frac{1}{2} n(z; \mu, \sigma^2 + \tau^2)} \\ &= \frac{n(z, \mu, \sigma^2 + \tau^2)}{n(z; \mu, \sigma^2) + n(z; \mu, \sigma^2 + \tau^2)} = \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2), \end{aligned} \quad (58)$$

where the latter is a notation we choose to define for simplicity. We plug (58) into (49), and that into (47) to get the final result

$$\begin{aligned} E[YI(X \neq Z)I(N = k + 1)r - c | Z = z, Y = 0 \cup X = Z \cup N > k] &= \\ \frac{p_Y \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2) (1 - p_N)^k p_N \cdot r}{(1 - p_Y) + p_Y \mathcal{N}_z^\mu(\sigma^2, \sigma^2 + \tau^2) + p_Y \mathcal{N}_z^\mu(\sigma^2 + \tau^2, \sigma^2) (1 - p_N)^k} - c. \end{aligned} \quad (59)$$