ECS647U / ECS773P
Bayesian Decision and Risk (BDRA)
Semester 6, 2021

Coursework 1 – Tricky Magicians

Instructions

This coursework requires an answer in three parts. Your answer should be prepared using a word processor and submitted electronically using QMPlus, as a single pdf file. All calculation/analysis steps must be shown. Use AgenaRisk for all calculations. Screen dumps of all results must be included along with specifications of expressions and probability values entered.

Also, where appropriate you should use simulation nodes in AgenaRisk.

It is strongly recommended that you consult the mandatory text for the course. In particular the beta binomial formulation.

This coursework is marked out of 100 and counts for 15% of the final mark for the module.

This is an individual coursework. ALL submissions will be carefully screened for signs of group work (including similar screen dumps and formatting) and if found College regulations governing assessment will apply.

Part 1: Background

1. Specify the Binomial distribution and describe under what circumstances its use is appropriate. **[4]**
2. Specify the probability density function for the Beta distribution and state how the expected value is calculated? **[4]**
3. What role do the parameters play in determining the shape of the distribution, and describe how it might be used to express a prior belief about the chance of a fair or unfair coin? **[6]**
4. Use a Beta-binomial formulation to specify the posterior distribution of P(*X* | *n*, *p*) where *X* is the number of heads, *n* is the number of flips and *p* is the probability of the coin coming up heads in each flip. **[6]**

Part 1: Answer [20]

1. $p(X=r)=\binom{n}{r}p^r(1-p)^{n-r}$. If r is the number of successes to occur in n repeated independent Bernoulli trials, each with probability of success p, then X is a Bernoulli random variable with parameters *n* and *p*.

2. $f(X)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} & \alpha>1,\beta>1 \\ 0 & \end{cases}$.

    Expected value: $E[X]=\dfrac{\alpha}{\alpha+\beta}$

3. Alpha represents successes and alpha+beta represents total number of trials when used as conjugate prior for binomial.

Prior belief for p, probability of coin coming up H/T, is expressed using Beta such that the expected value of p is: $E[X]=\dfrac{\alpha}{\alpha+\beta}$

4. $P(p\,|\,X,n)$ is calculated by Bayes theorem:

$$\begin{aligned}&P(p\,|\,X,n)=P(X\,|\,p,n)P(p)\\&=Binomial(X\,|\,p,n)Beta(\alpha,\beta)\end{aligned}$$

Part 2: Basic Analysis

Magicians often shave the edge of a coin to make it biased towards heads or tails when it is *spun*. The shaving is slight or shallow enough to remain undetectable by the naked eye. Shaving the head side makes the coin heavier on the tail side and therefore biased towards tails and vice versa.

You have been handed a coin that *has* indeed been shaved by a magician, but you are indifferent to which side has been shaved.

1. Model two hypothesised prior beliefs about the coin coming up heads on a given spin

   a) shaven on heads
   b) shaven on tails

   Use the Beta distribution where appropriate and assume that the strength of you belief is equivalent to a notional 30 coin spins. Also assume the hypothesised prior Beta distributions for p are modelled using parameter ratios of 20:10 and 10:20.

   Carefully describe and justify your model and show a graph of the prior distribution for p, as a probability density function or histogram, for the chance of heads on any future spin under each hypothesis. **[20]**

2. You are allowed 25 spins of the shaved coin and you observe the following results:

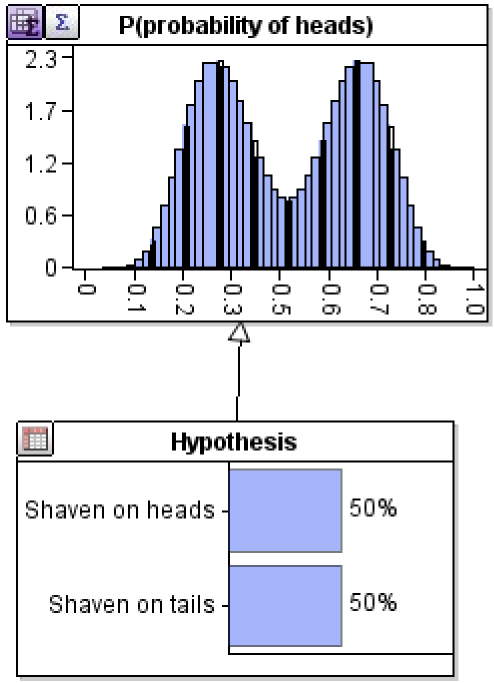
   H H H T H H T H H H T T H T H H T H H H H T H T H

   Carefully describe and justify how you used your model and calculate the posterior probability of each hypothesis. **[20]**

Part 2: Answer

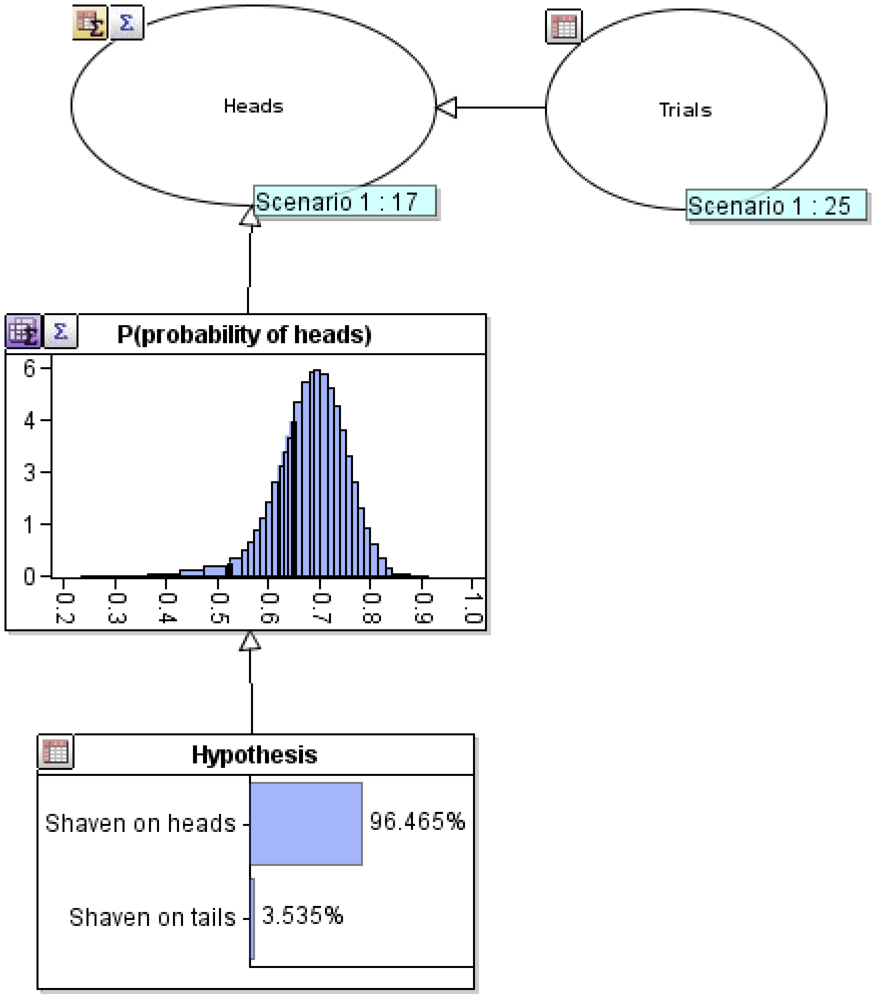
1. It should be clear from the answer that *alpha* + *beta* = 30. For a) shaven on heads the number of successes, alpha, equals the number of heads and should be 20. For b) shaven on tails the number of successes, alpha, equals the number of heads and should 10 since tails is more likely. Answers that assumed shaving on heads made tails more likely were not penalised.

The prior risk graphs should have p(probability heads) as the vertical axis and the horizontal axis should be "probability of heads" . For shaven on heads the majority of the probability mass should be to the "right" of the graph (greater belief in heads). And vice versa for shaven on tails. There should be little probability mass around ½ since a fair coin is unlikely.



The hypotheses are shaved on heads/shaved on tails so we are interested in the posterior distribution for p(hypothesis) NOT the posterior distribution for p(heads). However those who argued that there was some ambiguity in the question and focused on the posterior probability for *p* were not unduly penalised.

2. The posterior distribution, where 17 heads from 25 tosses gives this result. The posterior probability is in favour of shaved on heads with probability approx. 96%.


Heads
Trials
Scenario 1 : 17
Scenario 1 : 25
P(probability of heads)
6
4
3
1
0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Hypothesis
Shaven on heads
96.465%
Shaven on tails
3.535%

Part 3: Advanced Analysis

Repeat part 2 of the analysis but where you are told that there is a 2:1 chance that the magician has shaved the head of the coin.

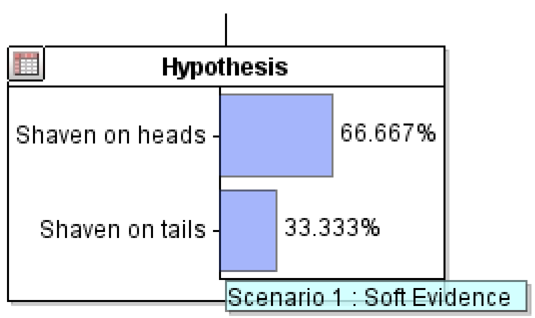
1. Describe how this information about the magician's bias would change your prior beliefs about each hypothesis. **[5]**
2. Carefully describe and justify your model and show a graph of the prior distribution for *p*, as a probability density function or histogram, for the chance of heads on any future spin under each hypothesis. **[7]**
3. Carefully describe and justify how you used your model and calculate the posterior probability of each hypothesis after having observed 25 coin spins: **[8]**

   H H H T H H T H H H T T H T H H T H H H H T H T H

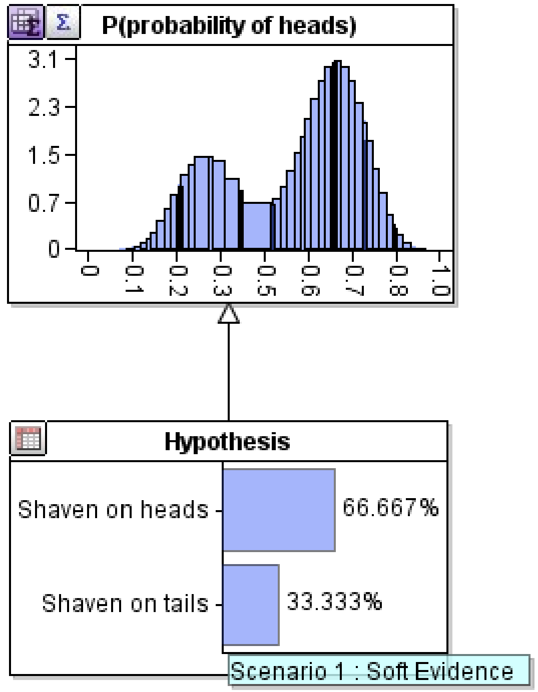
4. Rather than use the binomial distribution use a series of individual single Bernoulli trials in your model and explain how your model has changed to accommodate this contrasting approach and discuss the differences, or otherwise in the result. **[20]**

Part 3: Answer

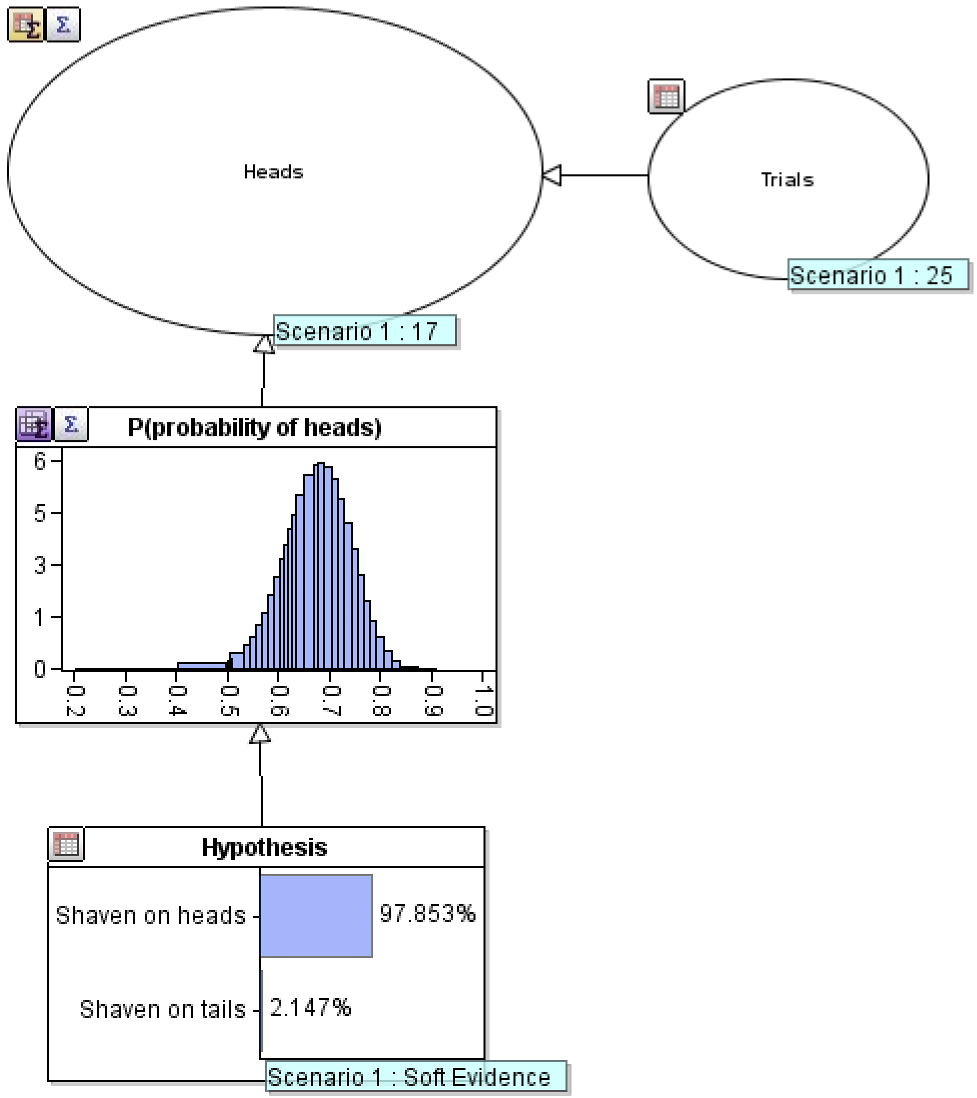
1. The prior belief in each hypothesis changes to p(shaved heads) = 66.67% p(shaved tails) = 33.33%. You can use soft evidence for this or change the NPT directly.



2. The posterior marginal distribution for p(heads) is shown below. This is multi-modal, but the key thing is that the probability mass should be concentrated above 0.5 because the coin is more likely to come up heads.



3. Now we enter the data from the coin spins:



New posterior on shaved heads hypothesis is 97.8%.

4. A Bernoulli trial is equivalent to a single Binomial trial with n = 1 and a fixed p for all trials. So in this model we would replace the binomial event node with 25 Bernoulli event nodes. The likelihood function would then be:

$$P(X \mid p, n) = Binomial(X \mid p, n) = \prod_{i=1}^{17} p^1(1-p)^0 \prod_{j=1}^{8} p^0(1-p)^1$$

The BN model is given below where 17 nodes have evidence of value one, for heads, and 8 nodes have value one, for tails. When comparing the single Binomial node solution and the multiple Bernoulli node solution the results are approximately equal.