MM925 Project

Introduction:

Echinococcus is a parasitic disease caused by the infection of a tape worm from the genus *Echinococcus* [1]. There are many animals such as dogs, foxes and small rodents that act as hosts for *Echinococcus* [2]. They are introduced to dogs by ingesting parasitic eggs in contaminated food and water, which develops into larval stage in the intestines [2]. There are two types of *Echinococcus*, one is cystic and the other is alveolar [1]. Alveolar *Echinococcus* is caused by the infection with larval stage of *Echinococcus multilocularis* (*E.multilocularis*) [1]. *E.multilocularis* can causes serious illnesses to humans [3]. *E.multilocularis* is characterised as asymptotic for 5-15 years [2]. Then slowly develops into a primary lesion located in the liver [2]. The symptoms include weight loss, bowel pain and fluid accumulation in the abdomen [3].

Currently, there are no known domestically acquired cases of *E.multilocularis* in Great Britain (GB) [3]. But this could not be said about other countries.

This report would like to assess the risk of the reintroduction of *E.multilocularis* back in to GB by dogs who have visited France with their owners on a 2 week holiday. It is allowed for the owner to take his dog provided the dog is treated for *E.multilocularis* before re-entering GB.

Method:

The mathematical model to estimate the probability of a dog becoming infected by *E.multilocularis* whilst on a holiday is given as,

$$p(t) = \frac{\beta(1-e^{-(\mu+\beta)t})}{\beta+\mu} \qquad (1)$$

where $p(t)$ is the probability of a dog becoming infected in t weeks, β is the infection pressure (transmission of infections over a period) in weeks in France and μ is the recovery rate in weeks. Using Equation 1 we derive a distribution which considers the uncertainty of β. Where β is between the range of 0.00001 and 0.0006 and most likely β would be 0.00018. It is also given with certainty that μ is 0.078. This derived distribution is the probability that a dog becomes infected during a 2-week holiday.

The distribution is derived from the understanding that β is uncertain and since we have no prior knowledge on how β behaves. The likelihood is modelled as an exponential distribution as there is a time frame and a rate at which the probability is estimated. This leads us to model β as a gamma distribution [4]. We then find the variability of the derived distribution of Equation 1 [5].

A second order model is created where the uncertainty is of β is given by the gamma distribution [6]. Thus, using the uncertainty of β into our Equation 1 we find the variability of the probability of a dog becoming infected in 2 weeks.

With the current compliance rules being that a dog travelling to France being treated in a time window of 24 to 120 hours when return to the GB. If this compliance rule is adhered to then a dog returning from a 2-week holiday to France would be considered infection free or treated. This suggests that if a the rules are not adhered to then the dog would be not treated. A survey was conducted in 2010, which found that 200 dog owners returned to GB and 20 of them did not treat their dogs. Using this survey, we have no prior knowledge therefore we model our prior probability as uniform distribution from 0 to 1, $(Uniform\sim(0,1))$ [4]. The likelihood would be modelled as binomial distribution of the number of dogs checked as 200 and 20 not being treated, $(Binomial\sim(200,20))$ [4]. Using Bayesian inference, the posterior probability would be modelled as beta distribution, since the number of dogs returning not treated is uncertain, $(Beta\sim(21,181))$ [4].

Since 2010, another survey had been conducted in 2022. This time 26 dog owners were surveyed and 1 dog had not been treated. We sequentially update our information using Bayesian inference. we know that the prior knowledge is the previous posterior probability distribution

$(Beta\sim(21,181))$ and the likelihood is same $(Binomial\sim(200,20))$ [7][10]. Thus, meaning that the new posterior probability would be modelled as a beta distribution $(Beta\sim(1,26))$ [7][10].

Of the 25 dogs in the survey that complied, their treatment times were taken. The data given follows a normal distribution, we use a parametric bootstrapping method to derive the distribution [8]. The uncertainty is within the mean and the standard deviation of the treatment times before entering GB therefore we ran a simulation for the 2 variables. The simulation ran for 1000 times and then the probability of the time being less than the 24 hours and the probability of the treatment time being greater than 120 had been summed up. This would allow us to derive the desired distribution. Using the compliance treatment time, we know that any time below 24 hours and above 120 hours would mean that the dog has not been treated.

Using the distributions from the Equation 1, the sequentially updated posterior distribution and finally the distribution of the treatment time, we derive a distribution which describes the probability of any dog returning to GB from France on a 2-week holiday is infected.
An equation is created by using the probability axioms [9]. We have 3 different events,

i) Event X: The dogs being infected,
ii) Event Y: The dogs not treated,
iii) Event Z: The dogs outside the treatment time.

The probability of these events are P(X), P(Y) and P(Z). The P(X) is independent since the dog would be infected or not as assumed in the Equation 1 [9]. The P(Y) is also independent since the owner could choose to neglect the compliance rules [9]. But the P(Z) is dependent on P(Y) since the dog owner adhered to the compliance rules of getting the dog treated and then we note the treatment time frame [9][10].
The leads us to create an equation for the dog not being treated,

$$P(Y) = P(Z|Y)P(Y) + P(Z'|Y)P(Y) \quad (2)$$

The assumption used for Equation 2 is that a dog not treated meant that the owner did not comply with the rules therefore the dog is not treated. It also does not make much sense that the being inside or outside the treatment time given they have not been treated.
Using Equation 2 we find that,

$$P(Any\ Dog\ is\ infected) = P(X) \times [P(Z|Y')P(Y') + P(Y)] \quad (3)$$

Using Equation 3 we derive a distribution of the probability of any dog returning to GB after a 2-week holiday is infected.

We assumed that 1000 dogs returned to GB from France on a 2 week holiday each year. We derive another distribution to describe the annual risk of the infection returning to GB. This is done by modifying the second order model previously stated, adjusting for the time, and running a loop for 1000 iterations. Then using Equation 3, instead of Event X being for just one dog it is now for 1000.

A scenario analysis is taken where we change the values of the parameter of β to 0.0003 and how this affects the annual risk of the infection returning to GB. We also modify the compliance rules with the time window being 10 to 72 hours and double the number of dogs returning from their holiday to 2000.
For all the derivations of distributions we will also summaries them with graphs and summary statistics for further numerical understanding.

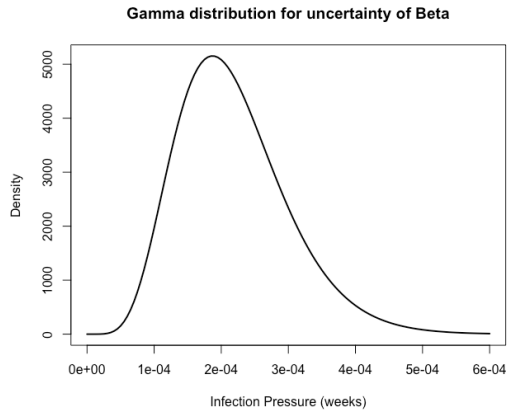

Figure 1: Gamma distribution of infection pressure, β, in weeks

Results:

Given β is between 0.00001 and 0.0006 and most likely 0.000187, we use trial and error to find the shape, a, and the scale, b, of the gamma distribution function [A1]. Figure 1 shows a gamma distribution of β which the data suggested with a peak at 0.000187. Using Figure 1 and Equation 1 in the second order model [A2], this with the uncertainty of β and the certainty of μ (0.078), we derive the distribution of the probability of a dog becoming infected during a 2-week holiday. We then simulate 10 different β and iterate 1000 different probabilities of dogs becoming infected [A3].

Figure 2 shows the empirical cumulative density function (ECDF) of the probability of a dog becoming infected during a 2-week holiday [A4]. This figure shows how the uncertainty of 10 different β effects Equation 1. The variability of Equation 1 is not large, the differences between the graphs are less than 0.0001. This is shown by the summary statistics of the different β [A5]. The first β produced used in Equation 1 produces a minimum of 5.99×10^{-5} and a maximum of 1.22×10^{-3}. The median is 3.84×10^{-4} and the confidence interval of 95% is $[1.69 \times 10^{-4}, 7.67 \times 10^{-4}]$. The second β produced used in Equation 1 has a minimum of 0.00009 and a maximum of 0.00107, The median is 3.95×10^{-4} and the confidence interval of 95% is $[1.68 \times 10^{-4}, 7.59 \times 10^{-4}]$. Comparing the two summary statistics we see that the difference between them is very small and therefore the variability of the two is also small. Which also then translates to the other 8 as seen in Figure 2. Visually the largest variation in the graph occurs when the probability of a dog infected is between 0.0004 and 0.0006.

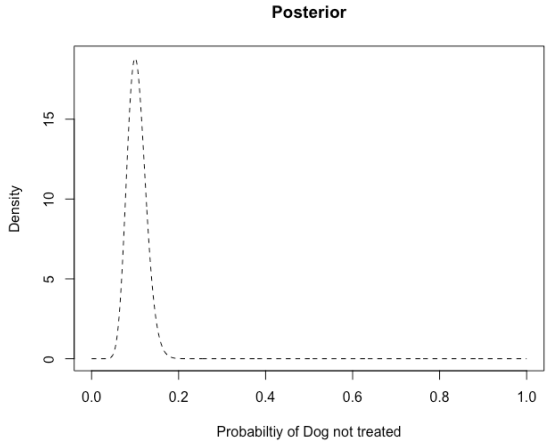

Figure 2: Empirical cumulative density function of the probability of a dog becoming infected during 2-week holiday.

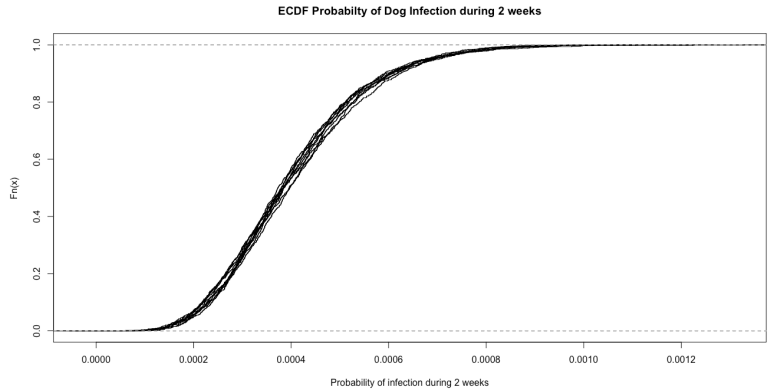

Figure 3: The posterior probability of a dog not treated in 2010 compliance rules.

The probability of a dog not being treated using the survey of the compliance rules is given in Figure 3 [A6]. The median of the posterior probability is 0.103 and the 95% confidence interval is $[0.0065, 0.149]$ [A7]. This means that the with 95% of the posterior probability of the dog not being treated is between 0.0065 and 0.149 also with the 50% of the posterior probability is 0.103.

Using Figure 4, we see that the likelihood and the prior probability are very similar [A8]. However, the posterior probability's peak has moved to the right closer to 0, but the range of the probability has increased compared to the prior probability. The median of the new posterior probability is 0.0613 and the 95% confidence interval is $[0.009, 0.189]$. This means that the 95% of the sequentially updated posterior probability has a larger range compared to the prior probability. But the median has become smaller[A9]. The impact of the prior to the posterior probability is that there is a larger variability in the posterior but a lower probability of the dog not being treated.

Figure 5 shows the density of the probability of the treatment time outside the 24 to 120 hours [A10]. The minimum is 0.033 and the maximum is 0.407. The median is 0.211 and the 95% confidence interval is $[0.094, 0.345]$, meaning that 95% of the probability of the treatment time outside 24 to 120 hours is between 0.094 and 0.345 [A11].

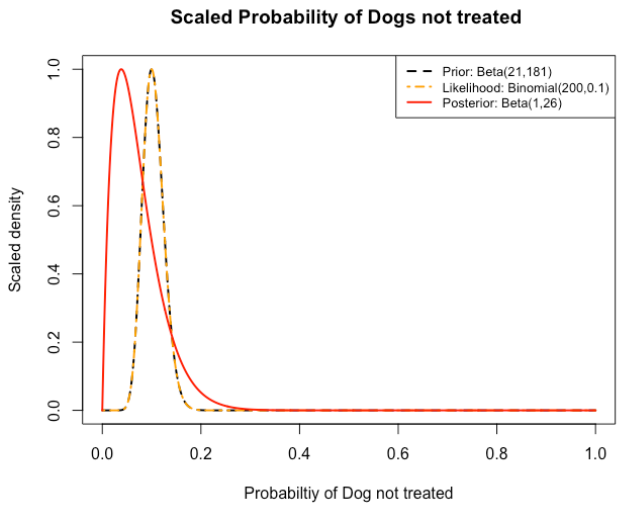

Figure 4: Scaled probability distribution of dogs not being treated.

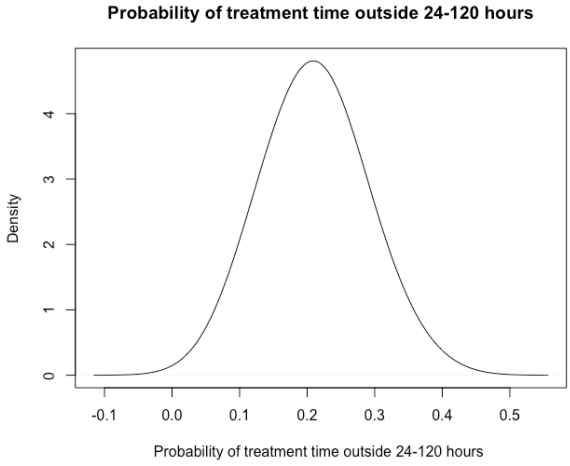

Figure 5: Density of the probability of treatment time outside 24-120 hours

Using the results from the distributions of Figures 2, 3 and 5 and Equation 3, we derive a distribution of the probability of any dog return to GB is infected. Figure 6 shows the probability of any dog returning to GB which is infected [A12]. The first probability produced a minimum of 8.88×10^{-6} and a maximum of 6.57×10^{-4}. The median is 8.67×10^{-5} and the 95% confidence interval is $[2.80 \times 10^{-5}, 2.89 \times 10^{-4}]$. The second probability produced a minimum of 1.34×10^{-5} and a maximum of 4.89×10^{-4}. The median is 8.59×10^{-5} and the 95% confidence interval is $[2.51 \times 10^{-5}, 2.77 \times 10^{-4}]$ [A13]. Comparing the two shows that the difference in-between then is not large meaning that the variability is also not large between the two probabilities. This is further shown in Figure 6 as the 10 different probabilities are very close despite the uncertainty on the β. There is a larger variation when the probability is between 1.5×10^{-4} and 2×10^{-4}. This is also noticed in Figure 2.

The annual risk probability of 1000 dogs infected during a 2-week holiday is shown in Figure 7 [A14]. This shows that minimum is 2.32×10^{-7} and a maximum of 1.10×10^{-5}. The median is 1.85×10^{-6} and the 95% confidence interval is $[5.68 \times 10^{-7}, 5.41 \times 10^{-6}]$ [A15] . This means that the annual probability risk is around 1.85×10^{-6} and 95% of the probability will be between $5.68 \times 10^{-7}$and 5.41×10^{-6}. The annual risk density has a similar shape to that of the Figure 1.

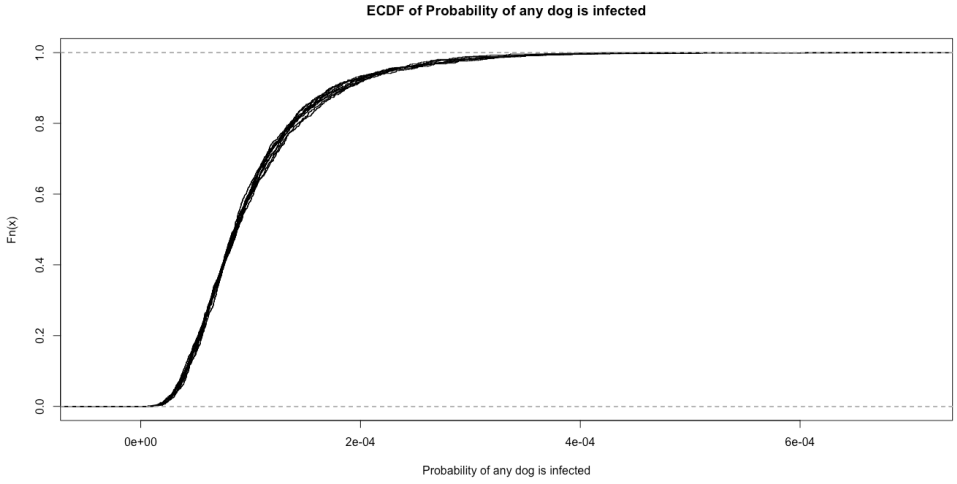

Figure 6: ECDF of the probability of any dog is infected returning to GB.

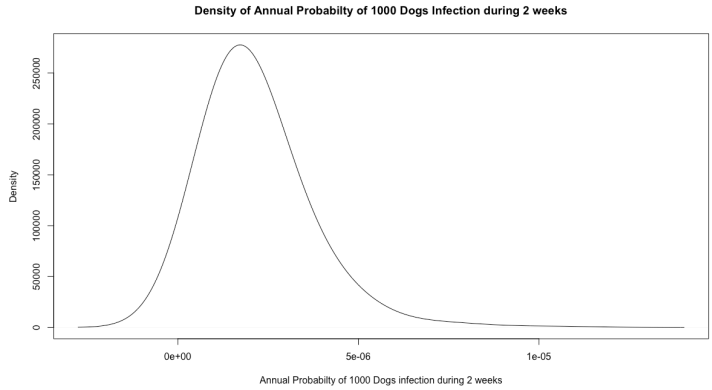

Figure 7: Density of the annual probability of 1000 dogs infected for 2 weeks.

A scenario analysis is undertaken and see the effects of increasing the β and the number of dogs, also decreasing the time window. Figure 8 plot 1 shows that the peak is now 0.0003 and the maximum has also increased to 0.0008. The effect of this is shown in plot 2 where the annual probability has also increased compared to Figure 7. The median is now 1.27×10^{-5} and the 95% confidence interval is [5.74×10^{-6}, 2.60×10^{-5}] [A16]. This means that an increase in β increase the annual probability.

The time window decreasing to 10 to 72 hours has increased the probability of a dog not being treated as shown in plot 3 and comparing to Figure 5. The median increase to 0.553 and the 95% confidence interval is [0.416,0.678] [A17].

The plot 4 shows the effect of decreasing the time window increases the annual probability compared to Figure 7. The median is 4.64×10^{-6} and the 95% confidence interval is [1.90×10^{-6}, 9.60×10^{-6}] [A18]. This shows reduction in the treatment time window results in the annual probability.

Finally, an increase in the number of dogs to 2000, increases the annual probability but not by a large margin. Comparing plot 5 and Figure 7 they show a very similar plot. The median is now 1.13×10^{-7} and the 95% confidence interval is [5.35×10^{-7}, 6.00×10^{-6}] [A19].

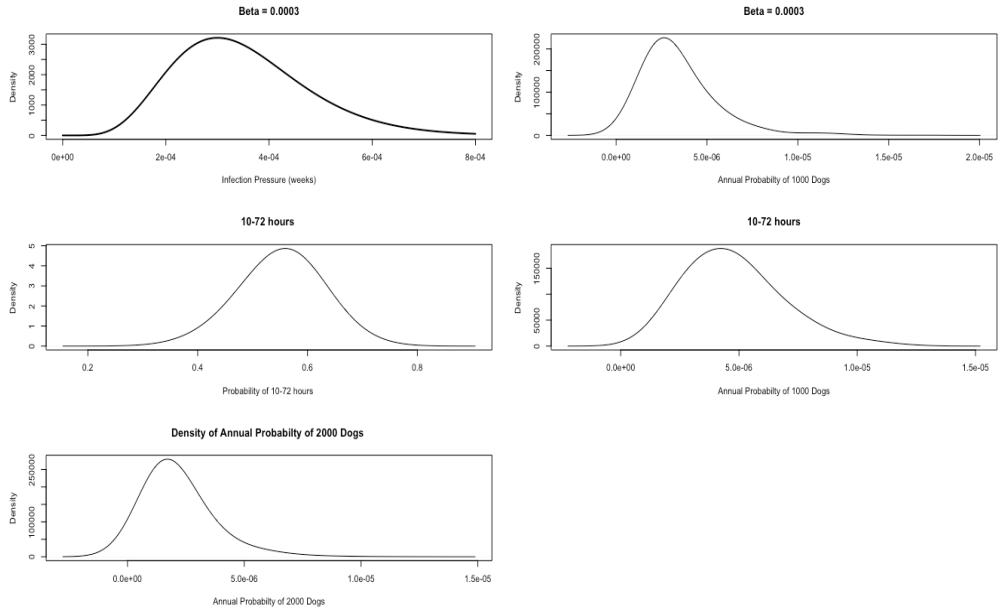

Figure 8: The density of infection pressure when $\beta = 0.0003$ (plot 1), the annual probability when $\beta = 0.0003$ (plot 2), probability of the 10-72 hours' time window (plot 3), the annual probability of the 10–72-hour time window (plot 4), The annual probability of 2000 dogs (plot 5)

Discussion:

In summary, we found that the uncertainty of infection pressure produces a small variation in the probability of a dog being infected during a 2-week holiday. The uncertainty in the mean and standard deviations from the normal distribution effects the compliance rules for the treatment times window of 24 to 120 hours.

The prior probability of the 2010 survey had a larger median and smaller confidence interval compared to the sequentially updated posterior probability of the 2022 survey.

The probability of any dog returning to the GB from their holiday is infected by *E.multilocularis* was very small given the uncertainty of the infection pressure.

These assumptions made for the derivation of the distribution could have a large effect. Especially the assumption that the probability of the event of the dog being infected is independent to the dog owner treating the dog or not. The assumption made of that the treatment time window was dependent on the dog owner treating the dog and adhering to the compliance rules.

The annual risk of the *E.multilocularis* introduction to GB was shown to be very small. This however depended on several factors, which were the length of the holiday to France, the dog owner not adhering to compliance rules by outright not treating the dog or not treating the dog within the period needed, also the number of dogs that travel in a year for 2 weeks.

The strength of the reports is the scenario analysis as this provides insight on how the different variables would affect the annual probability and by how much. The infection pressure increasing and the treatment time window decreasing had a largest effect which increased the probability of the infection *E.multilocularis* being introduced to GB. With a small increase from the number of dogs travelling annually.

The limitations of the report could be that assumptions of the probability axiom used in Equation 3 as this could affect the annual probability of the infection *E.multilocularis* being introduced to GB.

If the owner of the dog adheres to the treatment rules then the risk of *E.multilocularis* is very small and would have little risk to the public health.

Conclusion:

In conclusion, the risk of *E.multilocularis* being introduced to GB from France with dogs being the carrier of the infection is very small. The annual probability risk has a median of 1.85×10^{-6} with 95% confidence interval of the annual probability will be between 5.68×10^{-7} and 5.41×10^{-6}.

This is with the assumption that the dog has been in France for 2 weeks, the dog not being treated with or without complaining to the rules. The probability of annual risk of the introduction of the infection increases with an increase in the uncertainty of the infection rate and the uncertainty of the treatment time of the dog. If these two factors are carefully looked at the risk of the reintroduction of *E.multilocularis* to GB would reduce.

Policy makers are advised from this report not to decrease the treatment time window from 24-120 hours before the dog owner returns to GB. Also, enforce stricter rules on dog owners who don't treat their dogs (not listening to the compliance rules). This would ensure the safety of the public in Great Britain and remaining *E.multilocularis* free [3].

Reference:

[1] Centers for disease control and prevention, 31/05/23, ***Parasites-Echinococcosis [Online]***, https://www.cdc.gov/parasites/echinococcosis/index.html , Last Accessed 9/11/23.
[2] World Health Organization, 17/05/23, ***Echinococcosis [Online],*** https://www.who.int/news-room/fact-sheets/detail/echinococcosis , Last Accessed 9/11/23.
[3]Department of Environment, Food & Rural Affairs and Animal and Plant Health Agency [UKGOV], 27/01/21, ***Echinococcus multilocularis: how to spot and report the disease [Online]***, https://www.gov.uk/guidance/echinococcus-multilocularis-how-to-spot-and-report-the-disease , Last Accessed 9/11/23.
[4] Kavanagh, K.K, 15/12/20, ***4.12: Conjugate priors***, Last Accessed 9/11/23.
[5] Kelly, L.K, 30/10/20, ***3.3: Introduction to variability and uncertainty,*** Last Accessed 9/11/23.
[6] Kelly, L.K, 30/10/20, **3.5: Structuring a risk analysis model**, Last Accessed 9/11/23.
[7] Kavanagh, K.K, 18/12/20, **4.10: Estimating posteriors 1**, Last Accessed 9/11/23.
[8] Kelly, L.K, 2/11/20, ***5.4: Example: parametric bootstrapping***, Last Accessed 9/11/23.
[9] Kelly, L.K, 20/8/21, ***3.5: Axioms and Rules***, Last Accessed 9/11/23.
[10] Kavanagh, K.K, 30/10/20, ***4.4: Bayes theorem and sequential updating***, Last Accessed 9/11/23.

Appendix:

A1: (Question 1)

```
beta_mode <- 0.000187
x.range <- seq(0,0.0006,0.000005)
calc_b_gamma <- function(a, m){
  b<- m/(a-1)
  print(b)
}
a<- 7
b<- calc_b_gamma(a,beta_mode)
plot(x.range,dgamma(x.range, shape = a, scale = b), type = "l", lwd= 2,
     main = "Gamma distribution for uncertainty of Beta", xlab = "Infection Pressure (weeks)",
     ylab= "Density")
```

A2: (Question 1)

```
mu<- 0.078
simulate_p <- function(n_uncert,n_var,t){
  p<- array(dim = c(n_uncert,n_var))
  for (i in 1:n_uncert) {
    beta <- rgamma(n_var,shape = a, scale = b)
    for (j in 1:n_var) {
      p[i,j] <- (beta[j]*(1-exp(-(mu+beta[j])*t)))/(beta[j]+ mu)
    }
  }
  p
}
```

A3: (Question 1)

```
sim_p<- simulate_p(10,1000,2)
```

A4: (Question 1)

```
plot(ecdf(sim_p[1,]),do.p=FALSE,verticals=TRUE,xlab="Probability of infection during 2 weeks",
    main="ECDF Probabilty of Dog Infection during 2 weeks")
for (i in 2:10) {
 lines(ecdf(sim_p[i,]),do.p=FALSE,verticals=TRUE)
}
```

A5: (Question 1)

```
for(i in 1:10){
 new <- quantile(sim_p[i,], c(0,0.025,0.5,0.975,1))
 print(new)
}
```

```
            0%       2.5%            50%          97.5%       100%
5.987779e-05 1.694129e-04 3.839438e-04 7.670840e-04 1.226840e-03
            0%       2.5%            50%          97.5%       100%
0.0000936303 0.0001627251 0.0003800836 0.0007566432 0.0010651828
            0%       2.5%            50%          97.5%       100%
9.816333e-05 1.791647e-04 3.957155e-04 7.473075e-04 1.212600e-03
            0%       2.5%            50%          97.5%       100%
7.732707e-05 1.582593e-04 3.797614e-04 7.584069e-04 9.428154e-04
            0%       2.5%            50%          97.5%       100%
6.251861e-05 1.544893e-04 3.796661e-04 7.552048e-04 1.023652e-03
            0%       2.5%            50%          97.5%       100%
9.984556e-05 1.769128e-04 3.843005e-04 7.326806e-04 1.070165e-03
            0%       2.5%            50%          97.5%       100%
0.0000826240 0.0001613466 0.0003974547 0.0007710313 0.0014522321
            0%       2.5%            50%          97.5%       100%
0.0001134158 0.0001651278 0.0003833310 0.0007400070 0.0011100460
            0%       2.5%            50%          97.5%       100%
9.700574e-05 1.537895e-04 3.730418e-04 7.620373e-04 9.678368e-04
            0%       2.5%            50%          97.5%       100%
8.511553e-05 1.631212e-04 3.852496e-04 7.471006e-04 1.328636e-03
```

A6: (Question 2)

```
x <- seq(0,1,0.001)
n= 200
s= 20
posterior <- dbeta(x, s+1,n-s+1)
plot(x, posterior, main = "Posterior",xlab=" Probabiltiy of Dog not treated ",
    type="l", ylab="Density", col="black", lty=2)
```

A7: (Question 2)

```
lower_interval_old<-qbeta(0.025,s+1,n-s+1)
upper_interval_old<-qbeta(0.975, s+1,n-s+1)
median_old<-qbeta(0.5, s+1,n-s+1)
```

A8: (Question 3)

```
new_prior <- dbeta(x, s+1,n-s+1)
likelihood <- dbinom(s,n,x)
new_s <- 1
new_n <- 26
new_posterior <- dbeta(x, new_s+1,new_n-new_s+1)
plot(x, new_prior/max(new_prior), xlab="Probabiltiy of Dog not treated",
    main= "Scaled Probability of Dogs not treated ",
    type="l", ylab="Scaled density", lwd=2, col="black", lty=2)
lines(x,likelihood/max(likelihood),lwd=2, col="orange", lty=4)
lines(x,new_posterior/max(new_posterior), lwd=2, col="red", lty=1)
legend("topright", c("Prior: Beta(21,181)",
                "Likelihood: Binomial(200,0.1)",
                "Posterior: Beta(1,26)"),
      col=c("black","orange", "red"), lty=c(2,4,1), cex=0.8, lwd=2)
```

A9: (Question 3)

```
lower_interval_new<-qbeta(0.025,new_s+1,new_n-new_s+1)
upper_interval_new<-qbeta(0.975, new_s+1,new_n-new_s+1)
median_new<-qbeta(0.5, new_s+1,new_n-new_s+1)
```

A10: (Question 4)

```
dogs_pcb <- c(102, 76, 143, 50, 103, 36, 132, 38, 118, 58,142, 86, 48, 32, 4,
           39, 32, 112, 73, 28, 73, 90, 61, 42, 71)

pcb <- function(x, b){
  a <- array(dim = c(b,2))
  z <- c()
  m<-mean(x)
  s <- sd(x)
  for(i in 1:b){
    z<- rnorm(length(x), m,s)
    a[i,1] <- mean(z)
    a[i,2] <- sd(z)

  }
  a
}
sim_pcb <- pcb(dogs_pcb,1000)
under_24 <- c()
over_120 <- c()
outside_24_120 <- c()
for (i in 1:1000) {
  under_24[i] <- pnorm(24,sim_pcb[i,1],sim_pcb[i,2])
  over_120[i] <- 1-pnorm(120,sim_pcb[i,1],sim_pcb[i,2])
  outside_24_120[i] <- under_24[i]+over_120[i]
}
plot(density(outside_24_120, bw=0.05),  xlab = "Probability of treatment time outside 24-120 hours
", main= " Probability of treatment time outside 24-120 hours ")
```

A11: (Question 4)

```
quantile(outside_24_120, c(0,0.025 ,0.25,0.5,0.75,0.975,1))
          0%      2.5%       25%       50%       75%     97.5%      100%
0.03372832 0.09421405 0.16500902 0.21080443 0.25347700 0.34530929 0.40686693
```

A12: (Question 5)

```
treated_outside <- outside_24_120
not_treated <- new_posterior/max(posterior)
dog_infection <- sim_p

any_dog_is_infected <- (treated_outside*(1-not_treated)+ not_treated)*dog_infection

#Plot of ecdf
plot(ecdf(any_dog_is_infected[1,]),do.p=FALSE,verticals=TRUE,xlab="Probability of any dog is
infected",
    main="ECDF of Probability of any dog is infected")
for (i in 2:10) {
 lines(ecdf(any_dog_is_infected[i,]),do.p=FALSE,verticals=TRUE)
}
```

A13: (Question 5)

```
for (i in 1:10) {
 summarystat <-quantile(any_dog_is_infected[i,], c(0,0.025 ,0.25,0.5,0.75,0.975,1))
 print(summarystat)
}
          0%       2.5%        25%        50%        75%      97.5%       100%
8.883691e-06 2.800985e-05 5.901985e-05 8.665417e-05 1.269201e-04 2.886364e-04 6.568251e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
1.335943e-05 2.510830e-05 5.882081e-05 8.586599e-05 1.255862e-04 2.773894e-04 4.885470e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
9.905073e-06 2.768134e-05 6.112849e-05 8.778451e-05 1.299344e-04 2.791250e-04 4.545019e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
1.095735e-05 2.476432e-05 5.976665e-05 8.695414e-05 1.247511e-04 2.822494e-04 6.007516e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
1.088239e-05 2.517739e-05 5.766036e-05 8.580477e-05 1.249989e-04 2.602601e-04 6.050039e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
6.271058e-06 2.988368e-05 5.866047e-05 8.673527e-05 1.249693e-04 2.843447e-04 6.320571e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
6.977546e-06 2.691231e-05 5.908362e-05 8.759064e-05 1.311123e-04 2.883476e-04 9.575817e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
7.033834e-06 2.779304e-05 5.877423e-05 8.695417e-05 1.278699e-04 2.692345e-04 6.499099e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
1.057048e-05 2.617219e-05 5.808036e-05 8.494874e-05 1.211078e-04 2.634206e-04 6.062701e-04
          0%       2.5%        25%        50%        75%      97.5%       100%
6.989695e-06 2.703880e-05 5.922821e-05 8.800769e-05 1.249961e-04 2.653757e-04 4.194013e-04
```

A14: (Question 6)

```
simulate_p_yearly <- function(n_uncert,n_var,t){
  p<- array(dim = c(n_uncert,n_var))
  for (i in 1:n_uncert) {
    beta <- rgamma(n_var,shape = a, scale = b)
    for (j in 1:n_var) {
      p[i,j] <- (beta[j]*(1-exp(-(mu+beta[j])*(t/52))))/(beta[j]+ mu)
    }
  }
  p
}
#simulate p
num_of_dogs <- 1000
sim_p_yearly <- c()
for(i in 1:num_of_dogs){
  sim_p_yearly[i]<- simulate_p_yearly(10,1000,2)
}

n_dog_is_infected <- sim_p_yearly
n_dog_is_infected_yearly<-(treated_outside*(1-not_treated)+ not_treated)*n_dog_is_infected

plot(ecdf(n_dog_is_infected_yearly),do.p=FALSE,verticals=TRUE,xlab="Annual Probabilty of 1000
Dogs infection during 2 weeks",
     main="ECDF of Annual Probabilty of 1000 Dogs Infection during 2 weeks")
```

A15: (Question 6)

```
quantile(n_dog_is_infected, c(0,0.025 ,0.25,0.5,0.75,0.975,1))
        0%          2.5%           25%           50%        75%          97.5%          100%
2.322486e-07 5.682199e-07 1.228832e-06 1.851064e-06 2.655110e-06 5.408348e-06 1.102921e-05
```

A16: (Question 7)

```
beta_mode_new <- 0.0003
x.range <- seq(0,0.0008,0.000005)

calc_b_gamma <- function(a, m){
  b<- m/(a-1)
  print(b)
}

a_new<- 7
b_new<- calc_b_gamma(a,beta_mode_new)

plot(x.range,dgamma(x.range, shape = a_new, scale = b_new), type = "l", lwd= 2,
     main = "Gamma distribution for uncertainty of Beta", xlab = "Infection Pressure (weeks)",
     ylab= "Density")

simulate_p_yearly_new <- function(n_uncert,n_var,t){
  p<- array(dim = c(n_uncert,n_var))
  for (i in 1:n_uncert) {
    beta <- rgamma(n_var,shape = a_new, scale = b_new)
    for (j in 1:n_var) {
      p[i,j] <- (beta[j]*(1-exp(-(mu+beta[j])*(t/52))))/(beta[j]+ mu)
    }
  }
  p
}
#simulate p
num_of_dogs <- 1000
sim_p_yearly_new <- c()
for(i in 1:num_of_dogs){
  sim_p_yearly_new[i]<- simulate_p_yearly_new(10,1000,2)
}

n_dog_is_infected_new <- sim_p_yearly_new
n_dog_is_infected_yearly_new<-(treated_outside*(1-not_treated)+
not_treated)*n_dog_is_infected_new

plot(density(n_dog_is_infected_yearly_new, bw= 0.000001),xlab="Annual Probabilty of 1000 Dogs
infection during 2 weeks",
     main="Density of Annual Probabilty of 1000 Dogs Infection during 2 weeks")

quantile(n_dog_is_infected_new, c(0,0.025 ,0.25,0.5,0.75,0.975,1))
            0%                2.5%         25%         50%                      75%       97.5%             100%
2.721476e-06 5.744686e-06 9.584094e-06 1.266741e-05 1.608088e-05 2.599216e-05 4.031647e-05
```

A17: (Question 7)

```
under_10 <- c()
over_72 <- c()
outside_10_72 <- c()
for (i in 1:1000) {
  under_10[i] <- pnorm(10,sim_pcb[i,1],sim_pcb[i,2])
  over_72[i] <- 1-pnorm(72,sim_pcb[i,1],sim_pcb[i,2])
  outside_10_72[i] <- under_10[i]+over_72[i]
}

plot(density(outside_10_72, bw=0.05),  xlab = "Probability of treatment time outside 24-120 hours ",
    main= " Probability of treatment time outside 10-72 hours ")
```

```
      0%     2.5%      25%      50%      75%    97.5%     100%
0.2991703 0.4163277 0.5117642 0.5536708 0.5950125 0.6784023 0.7642239
```

A18: (Question 7)

```
treated_outside_new <- outside_10_72

n_dog_is_infected_yearly_new_outside<-(treated_outside_new*(1-not_treated)+
not_treated)*n_dog_is_infected

plot(ecdf(n_dog_is_infected_yearly_new_outside),do.p=FALSE,verticals=TRUE,xlab="Annual
Probabilty of 1000 Dogs infection during 2 weeks",
    main="ECDF of Annual Probabilty of 1000 Dogs Infection during 2 weeks")
plot(density(n_dog_is_infected_yearly_new_outside, bw= 0.000001),xlab="Annual Probabilty of
1000 Dogs infection during 2 weeks",
    main="Density of Annual Probabilty of 1000 Dogs Infection during 2 weeks")

quantile(n_dog_is_infected_yearly_new_outside, c(0,0.025 ,0.25,0.5,0.75,0.975,1))
      0%           2.5%            25%            50%         75%           97.5%           100%
7.616459e-07 1.897218e-06 3.436808e-06 4.636860e-06 6.068605e-06 9.599027e-06 1.620027e-05
```

A19: (Question 7)

```
num_of_dogs_new <- 2000
sim_p_yearly_new <- c()
for(i in 1:num_of_dogs_new){
  sim_p_yearly_new[i]<- simulate_p_yearly(10,1000,2)
}

n_dog_is_infected_new_num_dog <- sim_p_yearly_new
n_dog_is_infected_yearly_new_num_dog<-(treated_outside*(1-not_treated)+
not_treated)*n_dog_is_infected_new_num_dog

plot(density(n_dog_is_infected_yearly_new_num_dog),xlab="Annual Probabilty of 1000 Dogs
infection during 2 weeks",
    main="Density of Annual Probabilty of 1000 Dogs Infection during 2 weeks")

quantile(n_dog_is_infected_yearly_new_num_dog, c(0,0.025 ,0.25,0.5,0.75,0.975,1))
       0%           2.5%            25%        50%                   75%       97.5%           100%
1.131474e-07 5.351067e-07 1.215016e-06 1.777735e-06 2.562456e-06 6.001045e-06 1.130345e-05
```