STAT432 Assignment 2

Rory Sarten 301005654

21 August 2020

Question 1

a)

These are independent binomial processes (despite happening in sequence), so the three trapping occasions can be denoted:

$$n_1 \sim Binomial(N, \theta) = \binom{N}{n_1} \theta^{n_1} (1-\theta)^{N-n_1}$$
$$n_2 \sim Binomial(N - n_1, \theta) = \binom{N - n_1}{n_2} \theta^{n_2} (1-\theta)^{N-n_1-n_2}$$
$$n_3 \sim Binomial(N - n_1 - n_2, \theta) = \binom{N - n_1 - n_2}{n_3} \theta^{n_3} (1-\theta)^{N-n_1-n_2-n_3}$$

So the probability of catching $n_1 + n_2 + n_3$ rats is given by

$$Binomial(N, \theta) \times Binomial(N - n_1, \theta) \times Binomial(N - n_1 - N - 2, \theta)$$

b)

$$\begin{aligned}\ell(N, \theta) = \ln[&\frac{N!}{n_1(N-n_1)!} \theta^{n_1} (1-\theta)^{N-n_1} \\ &\times \frac{(N-n_1)!}{n_2(N-n_1-n_2)!} \theta^{n_2} (1-\theta)^{N-n_1-n_2} \\ &\times \frac{(N-n_1-n_2)!}{n_3(N-n_1-n_2-n_3)!} \theta^{n_3} (1-\theta)^{N-n_1-n_2-n_3}]\end{aligned}$$

$$\begin{aligned}\ell(N, \theta) = &\ln N! - [\ln n_1 + \ln(N - n_1)] + n_1 \ln \theta + (N - n_1) \ln(1-\theta) + \\ &\ln(N-n_1)! - [\ln n_2 + \ln(N - n_1 - n_2)] + n_2 \ln \theta + (N - n_1 - n_2) \ln(1-\theta) + \\ &\ln(N-n_1-n_2)! - [\ln n_3 + \ln(N - n_1 - n_2 - n_3)] + n_3 \ln \theta + (N - n_1 - n_2 - n_3) \ln(1-\theta)\end{aligned}$$

We can reduce the components

$n_1 \ln \theta + n_2 \ln \theta + n_3 \ln \theta = (n_1 + n_2 + n_3) \ln \theta$

$(N - n_1) \ln(1-\theta) + (N - n_1 - n_2) \ln(1-\theta) + (N - n_1 - n_2 - n_3) \ln(1-\theta) = (3N - 3n_1 - 2n_2 - n_3) \ln(1-\theta)$

$\ln N! + \ln(N - n_1)!...$

Gives

$$\ln N! - \ln(N - n_1 - n_2 - n_3)! + (n_1 + n_2 + n_3) \ln \theta + (3N - 3n_1 - 2n_2 - n_3) \ln(1-\theta)$$

finish that bit of algebra somehow

c)

```r
llfunc <- function(par, n1, n2, n3) {
  N <- par[1]
  theta <- par[2]
  lfactorial(N) - lfactorial(N - n1 - n2 - n3) +
    sum(n1, n2, n3)*log(theta) +
    (3*N - 3*n1 - 2*n2 - n3)*log(1 - theta)
}

n1 <- 82L
n2 <- 54L
n3 <- 23L
## since N >= n1 + n2 + n3
N_0 <- n1 + n2 + n3
par_start <- c(N_0, 0.5)

optim_fit <- optim(par = par_start,
                   fn = llfunc,
                   n1 = n1, n2 = n2, n3 = n3,
                   lower = c(N_0, 1e-4),
                   upper = c(Inf, 1),
                   method = "L-BFGS-B",
                   control = list(fnscale = -1),
                   hessian = TRUE)

MLE <- optim_fit$par
SE <- sqrt(diag(solve(-optim_fit$hessian)))
LowerBound <- MLE - qnorm(0.975) * SE
UpperBound <- MLE + qnorm(0.975) * SE
results <- cbind(MLE, SE, LowerBound, UpperBound) %>% round(2)
rownames(results) <- c("N", "theta")
results
```

```
##          MLE    SE LowerBound UpperBound
## N     189.43 13.23     163.49     215.36
## theta   0.45  0.06       0.34       0.57
```

The maximum likelihood point $\hat{N} = 189.43$ with confidence interval $(163.49, 215.36)$

The maximum likelihood point $\hat{\theta} = 0.45$ with confidence interval $(0.34, 0.57)$

d)

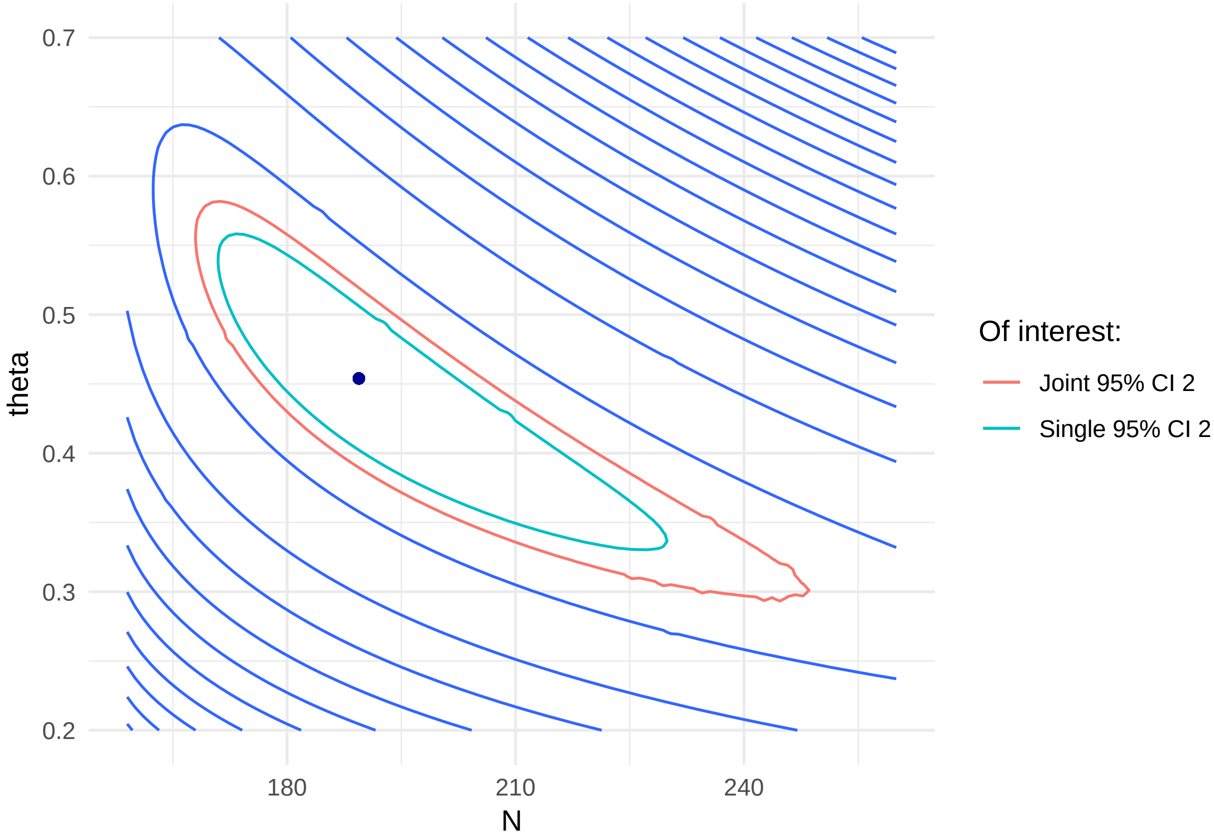
Contour map of (N, θ) surface

```r
n_vals <- seq(from = n1+n2+n3, to = 260, length = 100)
theta_vals <- seq(from = 0.2, to = 0.7, length = 100)

combos <- expand.grid(n_vals, theta_vals)
surface <- apply(expand.grid(n_vals, theta_vals), 1, llfunc, n1 = n1, n2 = n2, n3 = n3) %>% round(2)
outcome <- cbind(combos, surface)
names(outcome) <- c("N", "theta", "value")

ggplot(data = outcome, mapping = aes(N, theta)) +
  geom_contour(aes(z = value),
                   breaks = seq(from = min(outcome$value), to = max(outcome$value), by = 12)) +
```

```r
geom_contour(aes(z = value,
                 colour = factor(..level.. == max(surface) - 1.92,
                                 levels = c(F, T),
                                 labels = c("Single 95% CI "))),
             breaks = max(surface) - 1.92) +
geom_contour(aes(z = value,
                 colour = factor(..level.. == max(surface) - 3,
                                 levels = c(F, T),
                                 labels = c("Joint 95% CI "))),
             breaks = max(surface) - 3) +
geom_point(mapping = aes(x = MLE[1], y = MLE[2]), colour = "darkblue") +
labs(colour = "Of interest:") +
#ggtitle("Contour map of (N, \u03B8) surface") +
#ylab("\u03B8") +
theme_minimal()
```



e)

$$\frac{\partial \ell}{\partial \theta} = \frac{n_1 + n_2 + n_3}{\theta} - \frac{3N - 3n_1 - 2n_2 - n_3}{1 - \theta}$$

$$= \frac{(n_1 + n_2 + n_3) - (n_1 + n_2 + n_3)\theta - (3N - 3n_1 - 2n_2 - n_3)\theta}{\theta(1 - \theta)}$$

Setting this to 0

$$(n_1 + n_2 + n_3) - [(n_1 + n_2 + n_3) + (3N - 3n_1 - 2n_2 - n_3)]\theta = 0$$

$$[(n_1 + n_2 + n_3) + (3N - 3n_1 - 2n_2 - n_3)]\theta = (n_1 + n_2 + n_3)$$

$$\theta = \frac{(n_1 + n_2 + n_3)}{(n_1 + n_2 + n_3) + (3N - 3n_1 - 2n_2 - n_3)} = \frac{(n_1 + n_2 + n_3)}{3N - 3n_1 + n_1 - 2n_2 + n_2 - n_3 + n_3}$$

$$\hat{\theta} = \frac{n_1 + n_2 + n_3}{3N - 2n_1 - n_2}$$

f)

We have

$$\hat{\theta} = \frac{n_1 + n_2 + n_3}{3N - 2n_1 - n_2}$$

and

$$1 - \hat{\theta} = \frac{3N - 2n_1 - n_2}{3N - 2n_1 - n_2} - \frac{n_1 + n_2 + n_3}{3N - 2n_1 - n_2} = \frac{3N - 3n_1 - 2n_2 - n_3}{3N - 2n_1 - n_2}$$