732A66 Decision Theory

Assignment 1

Johannes Hedström

STIMA
Department of Computer and Information Science
Linköpings universitet

2024-10-07

Contents

```r
library(knitr)
```

1 Task 1

A bank official is concerned about the rate at which the bank's tellers provide service for their customers. He feels that all of the tellers work at about the same speed, which is either 30, 40 or 50 customers per hour. Furthermore, 40 customers per hour is twice as likely as each of the two other values, which are assumed to be equally likely. In order to obtain more information, the official observes all five tellers for a two-hour period, noting that 380 customers are served during that period. Show that the posterior probabilities for the three possible speeds are approximately 0.000045, 0.99996 and 0.00000012 respectively. (Hint: The total amount of 380 customers cannot be equalized with an average amount of customers per teller and hour. Using formal calculations instead of simulating data, it is easier to confirm the posterior probabilities.)

Prior probabilities:

$$P(30) = 0.25, P(40) = 0.5, P(50) = 0.25$$

$$r = 380, t = 2 \text{ hours times 5 tellers} = 10$$

We get the posterior probabilities from bayes theorem when making inference on λ in the poisson process:

$$P(\lambda \mid r, t) = \frac{\lambda^r e^{-t \cdot \ \lambda} \cdot P(\lambda)}{30^r e^{-t \cdot 30} \cdot 0.25 + 40^r e^{-t \cdot 40} \cdot 0.50 + 50^r e^{-t \cdot 50} \cdot 0.25}$$

Applying this on our data and lambda=30 would look like this:

$$P(\lambda = 30 \mid r = 380, t = 10) = \frac{30^{380} e^{-10 \cdot \ 30} \cdot 0.25}{30^{380} e^{-10 \cdot 30} \cdot 0.25 + 40^{380} e^{-10 \cdot 40} \cdot 0.50 + 50^{380} e^{-10 \cdot 50} \cdot 0.25}$$

This will create problems when calculating as 30^{380} will be infinity, so instead we take the log of the formula which gives us the following:

Likelihood * Prior:

$$log(\lambda^r e^{-t \cdot \ \lambda} \cdot 0.25) = r \cdot log(\lambda) - t\lambda - log(4)$$

The full expression is then:

$$\log(P(\lambda \mid r, t)) = r \log(\lambda) - t\lambda - \log(4) - \tag{1}$$
$$((r \ log(30) - 10 \cdot 30 - log(4)) + (r \ log(40) - 10 \cdot 40 - log(2)) + (r log(50) - 10 \cdot 50 - log(4))) \tag{2}$$

These calculations needs to be scaled to be able to get the exponent for the logged values, this can be done by subtracting with the largest likelihood as calculating exp on large negative values is not a computational problem.

This results in :

$$L(30, 480, 10) \cdot P(30) = 380 \cdot log(30) - 10 \cdot 30 - log(4) = 991$$

$$L(40, 480, 10) \cdot P(40) = r \cdot log(40) - 10 \cdot 40 - log(2) = 1001$$

$$L(50, 480, 10) \cdot P(50) = r \cdot log(\lambda) - 10 \cdot 50 - log(4) = 985$$

When subtracting with max we get:

$$LP_{30} - Max(LP) = 991 - 1001 = -10$$

$$LP_{40} - Max(LP) = 1001 - 1001 = 0$$

$$LP_{50} - Max(LP) = 985 - 1001 = -16$$

We can now anti log the values.

$$exp(-10) \approx 4.484337e - 05$$
$$exp(0) = 1$$
$$exp(-16) \approx 1.245451e - 07$$

Now we normalize these values by the sum of them:

$$P(\lambda = 30 \mid r = 380, t = 10) = \frac{4.484337e - 05}{4.484337e - 05 + 1 + 1.245451e - 07} \approx 0.000045$$
$$P(\lambda = 40 \mid r = 380, t = 10) = \frac{1}{4.484337e - 05 + 1 + 1.245451e - 07} \approx 0.99996$$
$$P(\lambda = 50 \mid r = 380, t = 10) = \frac{1.245451e - 07}{4.484337e - 05 + 1 + 1.245451e - 07} \approx 0.00000012$$

Below we have a function which checks that this will return the correct results

```
log_P <- function(lambda, r, t,p) {
  # First part of the expression: r * log(lambda) - t * lambda - log(4)
  part1 <- r * log(lambda) - t * lambda + log(p)

  # Subracting with maximum of the log values to rescale values
  part1 <- exp(part1-max(part1))

  # Final result: Normalizing the values with sum
  log_P_value <- part1 / sum(part1)

  return(log_P_value)
```

```
}

lambda <- c(30,40,50)
r <- 380
t <- 10
p <- c(0.25,0.5,0.25)

# running the function
res <- data.frame(log_P(lambda, r,t,p))

colnames(res) <- 'Posterior probabilities'

rownames(res) <- c('P(30)', 'P(40)', 'P(50)')

knitr::kable(res, digits=10)
```

	Posterior probabilities
P(30)	4.48414e-05
P(40)	9.99955e-01
P(50)	1.24500e-07

Another way of solving this question could be with utilizing the exponential class of distributions.

$$f(x|\theta) = f(x|\lambda) = e^{(ln\lambda)\cdot x - lnx! - \lambda}$$

Where $A(\lambda) = ln\lambda \cdot t$,$B(x) = 380$, $-lnx! = -\Gamma(x+1) = C = -\Gamma(381)$(constant) and $D(\lambda) = -\lambda \cdot t$.

This results in :

$$f(x|30) = e^{(ln300)\cdot 380 - \Gamma 381 - 300}$$

$$f(x|40) = e^{(ln400)\cdot 380 - \Gamma 381 - 400}$$

$$f(x|50) = e^{(ln500)\cdot 380 - \Gamma 381 - 500}$$

Then from these results use your prior and likelihood to get the posterior probabilities.

$$P(\lambda = 30|data) = \frac{f(x|30)\cdot 0.25}{f(x|30)\cdot 0.25 + f(x|40)\cdot 0.50 + f(x|50)\cdot 0.25}$$

$$P(\lambda = 40|data) = \frac{f(x|40)\cdot 0.50}{f(x|30)\cdot 0.25 + f(x|40)\cdot 0.50 + f(x|50)\cdot 0.25}$$

$$P(\lambda = 50|data) = \frac{f(x|50)\cdot 0.25}{f(x|30)\cdot 0.25 + f(x|40)\cdot 0.50 + f(x|50)\cdot 0.25}$$

```r
# calculating likelihood
lik <- exp(log(t*lambda)*r-lfactorial(r+1)-t*lambda)

# posterior probabilities
posteriot<- lik*p/(sum(lik*p))

res <- data.frame(posteriot)

colnames(res) <- 'Posterior probabilities'

rownames(res) <- c('P(30)', 'P(40)', 'P(50)')

knitr::kable(res, digits=10)
```

	Posterior probabilities
P(30)	4.48414e-05
P(40)	9.99955e-01
P(50)	1.24500e-07

Same results as in the question and as the first method, but this is the preferred way to solve the question as its a much simpler.

2 Task 2

Assume you have decided to bet on a horse race, and that you have very little knowledge about the competing horses. You consider betting on Little Joe, and you see that the odds for this horse are 9 to 1 (i.e. odds against that the horse will win). You decide to look up some historical tracks on how Little Joe has performed recently and note that he has won in 2 of the last 10 races he competed in. You can assume that these races are fairly comparable with respect to the levels of his competitors.

- a) Using the above as your background information, what are your subjective odds for Little Joe?
- b) If you bet, you will obtain 9 times the money you have put in the bet. What is your subjective expected return from betting on Little Joe?

2.1 a)

$$P_E = P(E|I, n, E) = \frac{a + n_E}{a + b + n}$$

$$P_E = P(E|I, n, E) = \frac{1 + 2}{1 + 9 + 10} = \frac{3}{20} = 0.15$$

$$Odds_E = \frac{P_E}{1 - P_E}$$

$$Odds_E = \frac{0.15}{0.85} = 3/17 = 0.1764706$$

```r
Pe <- 3/20 # prob

Odds_pe <- Pe/(1-Pe) # odds

3/17 == Odds_pe # check if the same
```

```
## [1] TRUE
```

My subjective odds is that the odds against that Little Joe will win is 17 to 3 and the odds for Little Joe is then 0.1764706.

2.2 b)

Subjective expected return

$$P_E \cdot 9 + (1 - P_E) \cdot -1$$

Using our subjective probabilities gives us:

$$0.15 \cdot 9 + 0.85 \cdot -1 = 0.5$$

```r
# Expected Return

(9*0.15) - 0.85
```

```
## [1] 0.5
```

The subjective expected return from the bet is 0.5, which mean that on average you expect to make a 50% profit for every money spent on the bet.