

# Worksheet 01 Group 2

Andrea Staub

Emanuel Mauch  
Jan Hohenheim

Guillaume Morlet  
Sophie Haldemann

Holly Vuarnoz

```
library(tidyverse)
library(rjags)
library(coda)
library(bayesmeta)
library(pCalibrate)
```

```
## Warning in .recacheSubclasses(def@className, def, env): undefined subclass
## "ndiMatrix" of class "replValueSp"; definition not updated
```

## Exercise 5

Prove the conditional Bayes theorem:

$$P(A|B, I) = \frac{P(B|A, I)P(A|I)}{P(B|I)}$$

### Proof

This is a conditional version of Bayes' theorem, where  $I$  is an additional piece of information (see STA421 Script, Chapter 1). We start with the definition of conditional probability:

$$P(A|B, I) = \frac{P(A \cap B|I)}{P(B|I)}$$

We know that conditional probability also satisfies:

$$P(B|A, I) = \frac{P(A \cap B|I)}{P(A|I)}$$

Rearranging terms, we get:

$$P(A \cap B|I) = P(B|A, I) \cdot P(A|I)$$

Substituting back into our original equation, we get:

$$P(A|B, I) = \frac{P(B|A, I) \cdot P(A|I)}{P(B|I)}$$

This completes the proof of the conditional Bayes theorem. It shows how to update the probability of an event  $A$  given the occurrence of another event  $B$  and additional information  $I$ , using prior probabilities and the likelihood of  $B$  given  $A$  and  $I$ .

## Exercise 6

### Given:

Diagnostic test with sensitivity  $P[T^+|D^+] = 0.95$  and specificity  $P[T^-|D^-] = 0.94$  Disease prevalence  $P[D^+] = 0.001$

### Question:

What is  $P[D^-|T^+]$ ?

### Solution:

Bayes:

$$P[A|B] = \frac{P[B|A] * P[A]}{P[B]}$$

Therefore:

$$P[D^-|T^+] = \frac{P[T^+|D^-] * P[D^-]}{P[T^+]}$$

$$P[T^+] = P[T^+|D^+] * P[D^+] + P[T^+|D^-] * P[D^-]$$

$$P[T^+|D^-] = 1 - P[T^-|D^-]$$

$$P[D^-] = 1 - P[D^+]$$

R code for calculation:

```
sens <- 0.95
spec <- 0.94
prev <- 0.001
pt <- sens*prev+(1-spec)*(1-prev)
prob_dn_tp <- ((1-spec)*(1-prev))/pt
prob_dn_tp
```

```
## [1] 0.9843981
```

The  $P[D^-|T^+]$  is 0.98. Therefore the probability that someone tested positive actually is healthy given a prevalence of the disease of 0.001 is very high due to the relatively low prevalence.

## Exercise 7

### Setup

```
set.seed(44566)
mu <- 160
sigma <- 20
M <- 1000
X <- rnorm(M, mu, sigma)
```

1.

```
glue::glue("True Values")
glue::glue("----")
glue::glue("mean: {mu}")
glue::glue("standard deviation: {sigma}")
glue::glue("variance: {sigma^2}")
```

```

glue::glue("median: {mu}")
glue::glue("0.025, 0.5 and 0.975 quantiles:")
qnorm(c(0.025, 0.5, 0.975), mu, sigma)

```

```

## True Values
## ----
## mean: 160
## standard deviation: 20
## variance: 400
## median: 160
## 0.025, 0.5 and 0.975 quantiles:
## [1] 120.8007 160.0000 199.1993

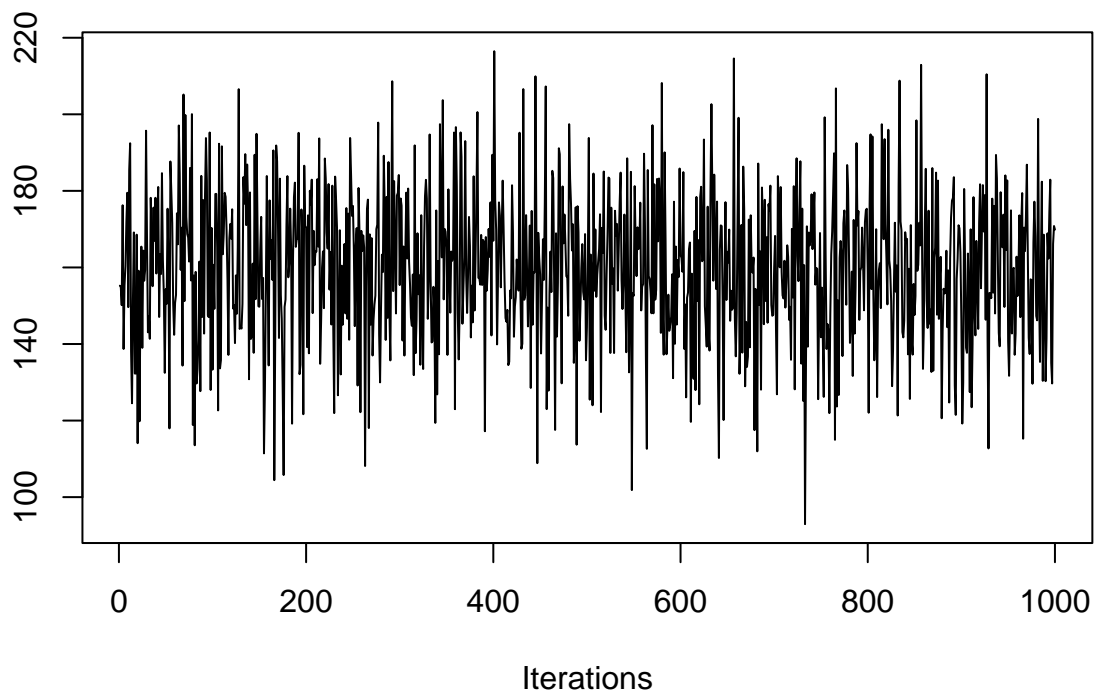
```

2.

```

coda::traceplot(mcmc(X))

```

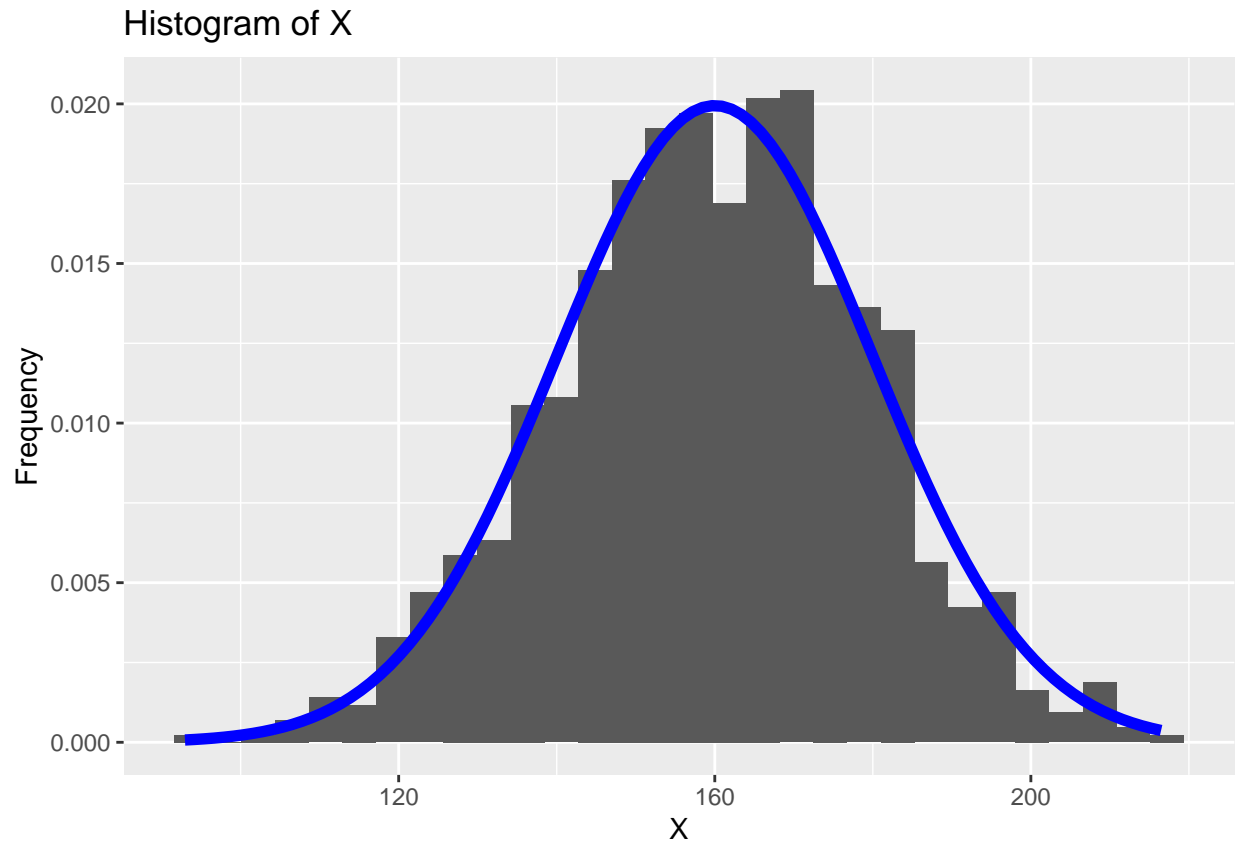


3.

```

tibble(x = X) |> ggplot(aes(x = x)) +
  geom_histogram(aes(y=after_stat(density)), bins = 30) +
  stat_function(fun = dnorm, args = list(mu, sigma), color="blue", linewidth = 2) +
  labs(title = "Histogram of X", x = "X", y = "Frequency")

```



The distributions are close, but the Monte Carlo histogram looks slightly bimodal.

4.

```
glue::glue("Estimated Values")
glue::glue("-----")
glue::glue("mean: {mean(X)}")
glue::glue("standard deviation: {sd(X)}")
glue::glue("variance: {var(X)}")
glue::glue("median: {median(X)}")
glue::glue("0.025, 0.5 and 0.975 quantiles:")
{quantile(X, probs = c(0.025, 0.5, 0.975))}
```

```
## Estimated Values
## -----
## mean: 159.612653448474
## standard deviation: 19.8154923956651
## variance: 392.653738882661
## median: 159.803466426713
## 0.025, 0.5 and 0.975 quantiles:
##      2.5%      50%      97.5%
## 119.6477 159.8035 197.4104
```

The estimations are fairly close to the truth

5.

```

glue::glue("Estimated P[X > 175]: {mean(X > 175)}")
glue::glue("True P[X > 175]: {pnorm(175, mu, sigma, lower.tail = FALSE)}")
glue::glue("")
glue::glue("Estimated P[150 < X < 180]: {mean(X > 150 & X < 180)}")
glue::glue("True P[150 < X < 180]: {pnorm(180, mu, sigma) - pnorm(150, mu, sigma)}")

## Estimated P[X > 175]: 0.225
## True P[X > 175]: 0.226627352376868
##
## Estimated P[150 < X < 180]: 0.536
## True P[150 < X < 180]: 0.532807207342556

```

The estimations are again pretty good

## Exercise 8

a

Likelihood:  $f(y|\mu) \sim N(\mu, \frac{1}{\kappa})$

Prior:  $f(\mu|H_1) \sim N(v, \frac{1}{\lambda})$

Marginal Likelihood:  $f(y|H_1) = \int f(y|\mu)f(\mu|H_1)d\mu$

The law of total expectation states that if  $X$  and  $Y$  are random variables on the same probability space, it is

$$E(X) = E(E(X|Y))$$

And similarly, the law of total variance states that

$$Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$$

(Held & Bové, p. 352).

Since the likelihood is normal with known variance, and the prior is also a normal distribution, we have a conjugate setting that determines that the marginal likelihood (also known as prior predictive) is normally distributed (Held & Bové, p. 234) with the following parameters:

$$Var(Y_{H_1}) = E[Var(Y_{H_1}|\mu_{H_1})] + Var[E(Y_{H_1}|\mu_{H_1})] = \frac{1}{\kappa} + \frac{1}{\lambda}$$

$$E(Y_{H_1}) = E[E(Y_{H_1}|\mu_{H_1})] = v$$

Plugging in the expectation  $v$  and the standard deviation  $\sqrt{\frac{1}{\kappa} + \frac{1}{\lambda}} = \sqrt{\frac{\kappa+\lambda}{\kappa\lambda}}$  into the PDF of a normal distribution, we get the desired result:

$$f(y|H_1) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\kappa\lambda}{\kappa+\lambda}} \exp \left\{ -\frac{\kappa\lambda}{2(\kappa+\lambda)} (y - v)^2 \right\}$$

b

The Bayes factor  $BF_{01}(y)$  is given by:

$$BF_{01}(y) = \frac{f(y|H_0)}{f(y|H_1)} = \frac{\sqrt{2\pi}\sqrt{\kappa} \exp \left( -\frac{\kappa}{2} (y - \mu_0)^2 \right)}{\sqrt{2\pi}\sqrt{\frac{\kappa\lambda}{\kappa+\lambda}} \exp \left( -\frac{\kappa\lambda}{2(\kappa+\lambda)} (y - v)^2 \right)}$$

After simplification, we get:

$$BF_{01}(y) = \sqrt{\frac{\kappa + \lambda}{\lambda}} \exp \left( -\frac{\kappa}{2}(y - \mu_0)^2 + \frac{\kappa\lambda}{2(\kappa + \lambda)}(y - v)^2 \right)$$

**c**

Examining the limits as  $\lambda$  approaches 0:

$$\lim_{\lambda \rightarrow 0} \left( \frac{\kappa + \lambda}{\lambda} \right) = \infty, \quad \lim_{\lambda \rightarrow 0} \left( \frac{\kappa\lambda}{2(\kappa + \lambda)} \right) = 0, \quad \text{therefore,} \quad \lim_{\lambda \rightarrow 0} BF_{01}(y) = \infty.$$

When lambda goes towards zero, the variance of the prior approaches infinity and therefore the prior will become an uninformative prior, therefore the Bayes factor is in favor of the null hypothesis and approaches infinity.

**d**

As established in 8a:  $Y|H_1 \sim N(v, \frac{1}{\kappa} + \frac{1}{\lambda})$ .

Now given  $\kappa = 1$ ,  $\lambda = 1/2$ , and  $v = 2$ :  $Y|H_1 \sim N(2, 3)$ .

The posterior probability  $P(H_0|y)$  can be calculated using Bayes' theorem (Held & Bové, p.232), where  $P(H_0)$  and  $P(H_1)$  cancel each other out since in our case they are equal:

$$\begin{aligned} P(H_0|y) &= \frac{f(y|H_0) \cdot P(H_0)}{f(y|H_0) \cdot P(H_0) + f(y|H_1) \cdot P(H_1)} \\ &= \frac{f(y|H_0)}{f(y|H_0) + f(y|H_1)} \\ &= \frac{f(y = 1|Y \sim \mathcal{N}(0, 1))}{f(y = 1|Y \sim \mathcal{N}(0, 1)) + f(y = 1|Y \sim \mathcal{N}(2, 3))} \\ &\approx 0.5538 \end{aligned}$$

Calculation in R:

```
Y_H0 <- dnorm(x = 1, mean = 0, sd = sqrt(1) )
Y_H1 <- dnorm(x = 1, mean = 2, sd = sqrt(3) )

HO_Y <- Y_H0 / (Y_H0 + Y_H1)
HO_Y
```

```
## [1] 0.5537843
```

The posterior of  $H_0$  is higher than its prior. This is because the two models  $H_0$  and  $H_1$  started out with equal priors, but the likelihood of seeing  $y=1$  is higher under  $H_0$  than under  $H_1$ .

Source: Held, L., & Bové, D. S. (2020). Likelihood and Bayesian inference: With applications in biology and medicine. Springer Berlin Heidelberg.

## Exercise 9

```
dat <- matrix(data = c(14, 9, 1, 5), ncol = 2, byrow = T)
bf <- twoby2Calibrate(dat, type = "two.sided", alternative = "simple")

formatBF(bf$minBF, digits = "default")
```

```
## [1] "1/3.4"
```

The bayes factor  $\frac{1}{3.4} = 0.29$  is decreasing the prior odds of no effect  $P[H_0]$ .

```
BF2pp(bf$minBF, prior.prob = 0.5)
```

```
## [1] 0.2270656
```

The posterior probability  $P[H_0|\text{data}]$  is estimated to be 0.23, meaning the probability that  $H_0$  (prior probability of no effect = 0.5) is true given our data is estimated to be 0.23.