

# Advanced Macroeconometrics – Assignment 2

Max Heinze ([h11742049@s.wu.ac.at](mailto:h11742049@s.wu.ac.at))      Gabriel Konecny ([h11775903@s.wu.ac.at](mailto:h11775903@s.wu.ac.at))  
Patrick Schüssele ([h11712262@s.wu.ac.at](mailto:h11712262@s.wu.ac.at))

May 10, 2023

## Contents

<b>Exercise 1</b>	<b>2</b>
Convergence of Means of $n$ Normal Draws . . . . .	2
Convergence of Means of $n$ Draws From a Cauchy Distribution . . . . .	2
<b>Exercise 2 (STILL TO DO)</b>	<b>4</b>
<b>Exercise 3</b>	<b>5</b>
Writing a Function to Simulate $\mathbf{y} = \alpha + \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ . . . . .	5
Simulating Data with $k = 1$ and $\sigma = 1$ . . . . .	5
Latent Values of the Model if we know $\sigma$ . . . . .	7
Potentially Interesting Regression . . . . .	7
Posterior Density for Simulated Data (STILL TO DO) . . . . .	7
<b>Exercise 4 (NEEDS REWRITING)</b>	<b>10</b>
Estimating $\mu$ . . . . .	10
Estimating $\sigma^2$ . . . . .	11
Including $\eta$ . . . . .	12

*The executable code that was used in compiling the assignment is available on GitHub at  
<https://github.com/maxmheinze/macrometrics>.*

## Exercise 1

### Convergence of Means of $n$ Normal Draws

We draw 100 times from a  $\mathcal{N}(5, 9)$  distribution and plot the cumulative mean as well as the expected value.

```
set.seed(2345)

n <- 100

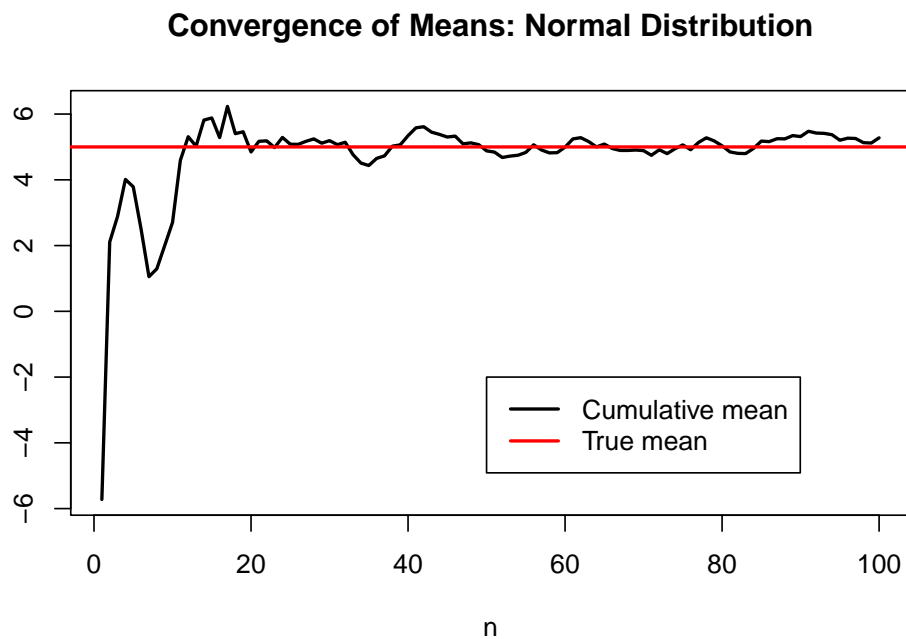
normal_data <- data.frame(n = 1:n, x = rnorm(n, mean = 5, sd = 9), mean = NA)

for (i in 1:n) {
  normal_data$mean[i] <- mean(normal_data$x[1:i])
}

plot(x = normal_data$n, y = normal_data$mean, type = "l", lwd = 2, main = "Convergence of
Means: Normal Distribution",
     xlab = "n", ylab = "")

abline(h = 5, col = "red", lwd = 2)

legend(50, -2, c("Cumulative mean", "True mean"), lwd = c(2, 2), col = c("black",
"red"))
```



We can easily see that the cumulative mean, that is,

$$\frac{1}{n} \sum_{i=1}^n x_i, \quad n = 1, \dots, 100,$$

converges to the true mean, the expected value of the distribution,  $E(N(5, 9)) = 5$ .

### Convergence of Means of $n$ Draws From a Cauchy Distribution

Next, we draw 10,000 times from a Cauchy distribution as specified and plot the cumulative mean.

```
set.seed(2345)

n <- 10000

cauchy_data <- data.frame(n = 1:n, x = rnorm(n, mean = 0, sd = 1)/rnorm(n, mean = 0,
```

```

sd = 1), mean = NA)

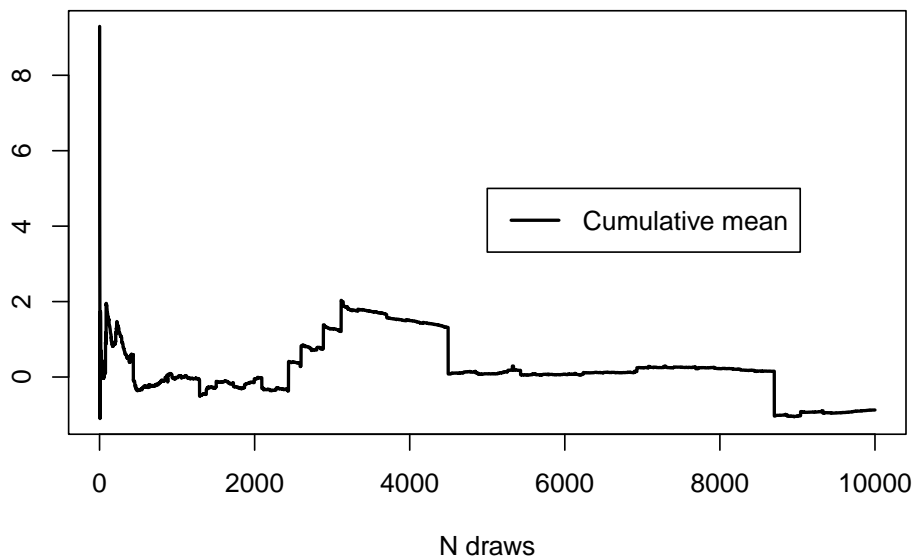
for (i in 1:n) {
  cauchy_data$mean[i] <- mean(cauchy_data$x[1:i])
}

plot(x = cauchy_data$n, y = cauchy_data$mean, type = "l", lwd = 2, main = "Convergence of
Means: Cauchy Distribution",
     xlab = "N draws", ylab = "")

legend(5000, 5, "Cumulative mean", lwd = 2, col = "black")

```

### Convergence of Means: Cauchy Distribution



This time, we do not observe convergence of the cumulative mean. Rather, we observe repeated random jumps in the cumulative mean. This is due to the Cauchy distribution not having an expected value that it could converge to.

## Exercise 2 (STILL TO DO)

## Exercise 3

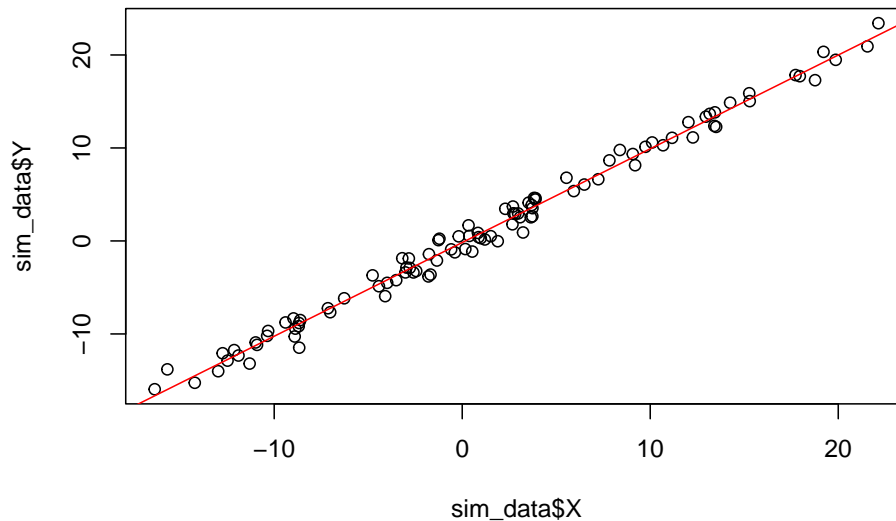
### Writing a Function to Simulate $y = \alpha + X\beta + e$

```
simulate_reg <- function(n, k = 1, alpha = 0, beta = 1, sigma = 1) {  
  
  # Remind the user to specify a vector as beta if they want to  
  if ((length(beta) == 1) * (k > 1) == 1) {  
    warning("You have specified k > 1 but only one value for beta.  
           This function then uses the same beta for all independent variables!")  
  }  
  
  # Create an empty matrix for the independent variables  
  X <- matrix(nrow = n, ncol = k)  
  
  # Fill them with draws from normal distributions. For fun we let the mean  
  # increase by 1 for every additional independent variable  
  for (i in 1:k) {  
    X[, i] <- rnorm(n, i, 10)  
  }  
  
  # Make beta a vector of length k if it is only a scalar  
  if ((length(beta) == 1) == TRUE) {  
    beta <- rep(beta, k)  
  }  
  
  # Compute  $y = a + Xb + e$ ,  $e \sim N(0, \sigma)$   
  simul_dep <- rep(alpha, n) + X %*% beta + rnorm(n, 0, sigma)  
  
  # Create one output data frame  
  output_data <- data.frame(Y = simul_dep, X)  
  
  # Return it  
  return(output_data)  
}
```

### Simulating Data with $k = 1$ and $\sigma = 1$

```
set.seed(2345)  
  
sim_data <- simulate_reg(100, 1, 0, 1, 1)  
  
plot(y = sim_data$Y, x = sim_data$X, main = "Simulated Data")  
  
abline(lm(sim_data$Y ~ sim_data$X), col = "red")
```

## Simulated Data



```
set.seed(2345)

coeffs <- vector(length = 1000)

for (i in 1:1000) {
  sim_data_2 <- simulate_reg(100, 1, 0, 1, 1)

  this_coefficient <- coef(lm(sim_data_2$Y ~ sim_data_2$X))[[2]]

  coeffs[i] <- this_coefficient
}

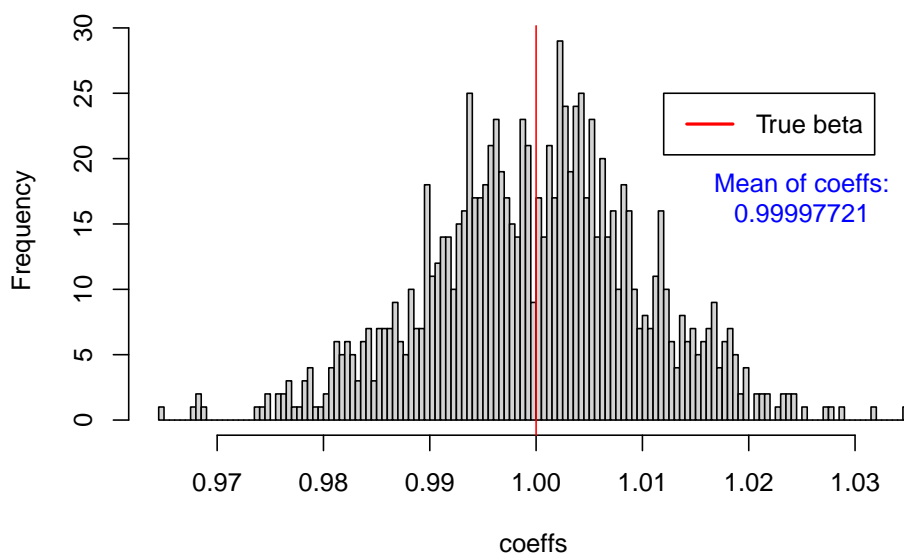
hist(coeffs, breaks = 100)

abline(v = 1, col = "red")

legend(1.012, 25, "True beta", lwd = 2, col = "red")

text(1.025, 17, paste0("Mean of coeffs:\n", round(mean(coeffs), 8)), col = "blue")
```

## Histogram of coeffs



The true value  $\beta_0 = 1$ . We see that if we run this regression over and over and over again, we get a distribution of estimated  $\beta$ s that is centered around the true value, but with some variance. Indeed, the mean of our estimated

betas equals 0.9999772, which is very close to  $\beta_0 = 1$ .

## Latent Values of the Model if we know $\sigma$

Let us say that we know  $\sigma$ . Then, since  $\sigma$  is known,  $\beta$  is the only latent parameter.

## Potentially Interesting Regression

Let's say we are interested in the determinants of our Econometrics grades, and we propose the model

$$\text{points}_i = \alpha + \text{study time}\beta_1 + \text{talent}\beta_2 + \text{likeability}\beta_3 + u_i,$$

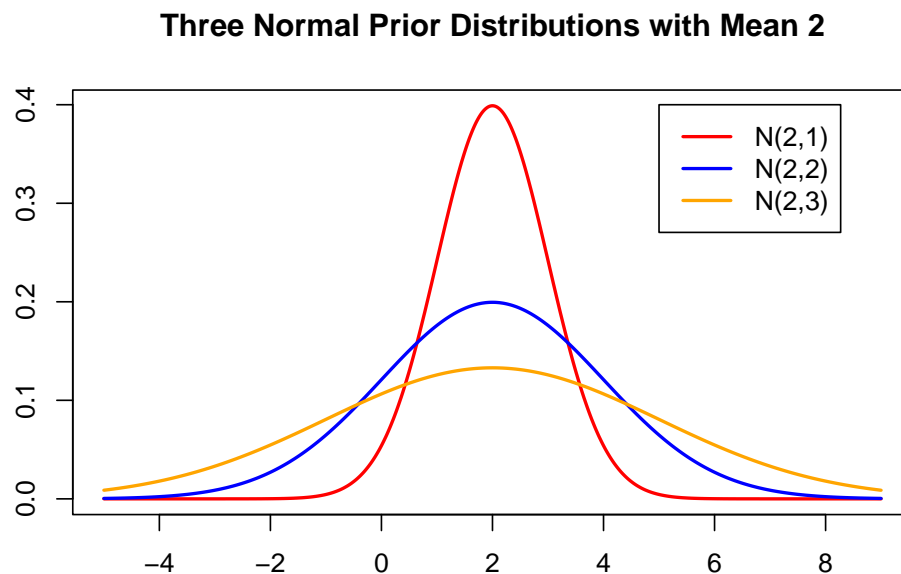
i.e. we regress the total course points (0–100) of student  $i$  on the time they devoted to studying, their inherent ability (assuming we can measure that), and how much they are deemed likable by the teachers of the course.

How can we come up with a prior distribution for  $\beta_1$ , the parameter associated with study time?

- It seems sensible to assume that the parameter associated with study time is **positive**, since it would be very weird if increasing the time devoted to studying for your Econometrics exam led to you getting a lower grade in your exam. If that were the case, nobody would be studying.
- We could even try to provide reasoning for a mean that we would want our prior distribution to be centered around. It would make sense if one hour of studying betters our result by, say, two points (and neither 0 nor 20). So we would use a prior distribution that is **centered at 2**.
- Then, we could assume for simplicity that the prior distribution is **normal**.

We could then go for one of the following priors:

```
plot(x = seq(-5, 9, length = 1000), y = dnorm(seq(-5, 9, length = 1000), 2, 1), type = "l",  
     lwd = 2, col = "red", main = "Three Normal Prior Distributions with Mean 2",  
     xlab = "", ylab = "")  
lines(x = seq(-5, 9, length = 1000), y = dnorm(seq(-5, 9, length = 1000), 2, 2),  
      type = "l", lwd = 2, col = "blue")  
lines(x = seq(-5, 9, length = 1000), y = dnorm(seq(-5, 9, length = 1000), 2, 3),  
      type = "l", lwd = 2, col = "orange")  
  
legend(5, 0.4, c("N(2,1)", "N(2,2)", "N(2,3)"), lwd = c(2, 2, 2), col = c("red",  
    "blue", "orange"))
```



## Posterior Density for Simulated Data (STILL TO DO)

```
simulate_reg(100, 1, 0, 1, 1)
```

##	Y	X
## 1	8.1106784	9.04651443
## 2	-5.0710065	-5.42795965
## 3	4.1455304	2.69556653
## 4	14.5061246	13.64135913
## 5	14.7413552	14.04097655
## 6	-4.0199372	-6.07476704
## 7	-8.2259197	-8.61974534
## 8	-1.1535692	0.86586999
## 9	10.3550099	8.19151734
## 10	3.2933633	2.04992850
## 11	4.4960863	4.39384975
## 12	2.9212602	2.13575171
## 13	-3.7429879	-4.49866429
## 14	19.5187429	19.48128641
## 15	4.5258026	4.78521069
## 16	-13.8998149	-13.95218432
## 17	21.0187735	21.20630653
## 18	20.0566884	20.47339103
## 19	-10.1505305	-9.69453435
## 20	1.1652733	1.33949196
## 21	-15.6540434	-17.31677512
## 22	-1.3672545	-0.64562066
## 23	-5.2541684	-4.69293570
## 24	13.2087607	13.60181982
## 25	-0.1155161	-0.04243287
## 26	0.9456359	0.22695943
## 27	2.8816332	2.81021147
## 28	6.0345664	6.04619499
## 29	11.4917864	12.03786633
## 30	5.9603796	4.71466749
## 31	10.5182950	12.68447782
## 32	5.0454010	5.23040514
## 33	1.4876452	-0.12116044
## 34	5.6775816	5.81664829
## 35	-9.2573947	-8.73366992
## 36	28.5614398	28.36827564
## 37	-9.0843589	-8.76435770
## 38	-2.4135167	-1.97933989
## 39	6.9078473	7.77491545
## 40	12.3165801	12.07459118
## 41	4.6256230	6.35577822
## 42	6.3875957	6.92182631
## 43	1.3735152	0.57878649
## 44	4.9890847	4.19963661
## 45	-14.6522498	-14.67825483
## 46	6.2842545	5.38396166
## 47	-0.6292212	-0.74128061
## 48	-12.0988964	-10.28112706
## 49	2.5853768	4.44049252
## 50	7.8111714	7.31735481
## 51	-2.0261854	-2.92863283
## 52	-9.9392815	-9.19870299
## 53	-5.2611799	-3.75054501
## 54	0.3946473	0.60286882
## 55	12.3535575	13.22441299
## 56	13.6806216	14.84516290
## 57	16.6861890	16.97409524
## 58	16.8058608	17.50741371
## 59	3.0529179	4.99672838
## 60	-3.4507261	-3.96907741



## 61	-15.4127941	-14.71794049
## 62	7.6644125	7.95946419
## 63	7.0886069	4.59741561
## 64	11.1098860	9.45538696
## 65	11.0641912	11.74136689
## 66	9.5980028	8.89954977
## 67	0.9673775	1.08563162
## 68	-13.5845008	-13.77922683
## 69	-6.0237998	-6.51521801
## 70	-0.4727045	-0.97410791
## 71	-6.9405263	-7.42466598
## 72	9.9396639	8.14506419
## 73	13.2323123	12.06384788
## 74	2.8993410	3.69531675
## 75	0.6645423	0.68631734
## 76	-5.0305368	-5.30934835
## 77	19.4683005	20.12256542
## 78	6.9859641	8.37377460
## 79	8.0715760	8.65619934
## 80	-2.3276217	-1.87776224
## 81	-20.8243709	-20.75277790
## 82	-4.2422050	-4.20300282
## 83	-2.8482576	-2.98605132
## 84	17.3156819	18.13895025
## 85	-10.4330064	-10.78191452
## 86	-15.6481156	-15.43257265
## 87	2.5443648	3.39586794
## 88	6.8444001	7.59759391
## 89	1.5921667	0.41754163
## 90	-1.7924879	-0.87917594
## 91	14.6087972	16.23012684
## 92	15.5140037	14.55581882
## 93	-1.6405400	-3.17225092
## 94	14.2410263	13.10479519
## 95	-1.4869144	-1.08802072
## 96	-1.1123803	-2.93732690
## 97	-4.3961368	-6.18058140
## 98	-1.7619484	-2.65182475
## 99	9.1761075	10.56098965
## 100	13.3092943	13.46167255

## Exercise 4 (NEEDS REWRITING)

### Estimating $\mu$

**Q:** Suppose you have data  $\mathbf{y} \sim N(\mu, 1)$ , and want to estimate  $\mu$ . Specify a Normal prior  $\mu \sim N(\mu_0, \sigma_0^2)$ . Derive the posterior  $p(\mu|\mathbf{y})$  by applying Bayes' theorem. Create histograms of two priors of your choice.

First, as we assume our data  $\mathbf{y} \sim N(\mu, 1)$ , the probability density of a single observation  $y_i$  is given by  $p(y_i|\mu, 1) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}(y_i - \mu)^2)$ . If we assume our data to be i.i.d., then the joint density (the likelihood function) is given by  $p(\mathbf{y}|\mu, 1) = \prod_{i=1}^n p(y_i|\mu, 1) = (2\pi)^{-\frac{n}{2}} \exp(-\frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2)$ .

We further assume a normal prior  $\mu \sim N(\mu_0, \sigma_0^2)$ , hence  $p(\mu|\mu_0, \sigma_0^2) = (2\pi\sigma_0^2)^{-\frac{1}{2}} \exp(-\frac{1}{2} \frac{(\mu - \mu_0)^2}{\sigma_0^2})$ .

To get to our posterior, we utilize Bayes' theorem, dropping all constant factors from our likelihood function and our prior in the process, since we are interested in proportionality:

$$\begin{aligned}
 p(\mu|\mathbf{y}, 1) &\propto p(\mathbf{y}|\mu, 1) \times p(\mu|\mu_0, \sigma_0^2) \\
 &\propto \exp(-\frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2) \times \exp(-\frac{1}{2} \frac{(\mu - \mu_0)^2}{\sigma_0^2}) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \mu)^2) + \frac{1}{\sigma_0^2} (\mu - \mu_0)^2) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n(\bar{y} - \mu)^2 + \frac{1}{\sigma_0^2} (\mu - \mu_0)^2) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n\bar{y}^2 - 2n\bar{y}\mu + n\mu^2 + \frac{1}{\sigma_0^2} \mu^2 - \frac{2}{\sigma_0^2} \mu\mu_0 + \frac{1}{\sigma_0^2} \mu_0^2) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n\bar{y}^2 + \frac{1}{\sigma_0^2} \mu_0^2 + (n + \frac{1}{\sigma_0^2}) \mu^2 - 2(n\bar{y} + \frac{1}{\sigma_0^2} \mu_0) \mu)
 \end{aligned}$$

Note that above, we define  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ , and utilized the fact that  $\sum_{i=1}^n (y_i - \mu)^2 = \sum_{i=1}^n (y_i - \bar{y} + \bar{y} - \mu)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2$ . Now let  $\mu_n = (n + \frac{1}{\sigma_0^2})^{-1} (n\bar{y} + \frac{1}{\sigma_0^2} \mu_0)$  and  $\sigma_n^2 = (n + \frac{1}{\sigma_0^2})^{-1}$ . Then we can further simplify the posterior to:

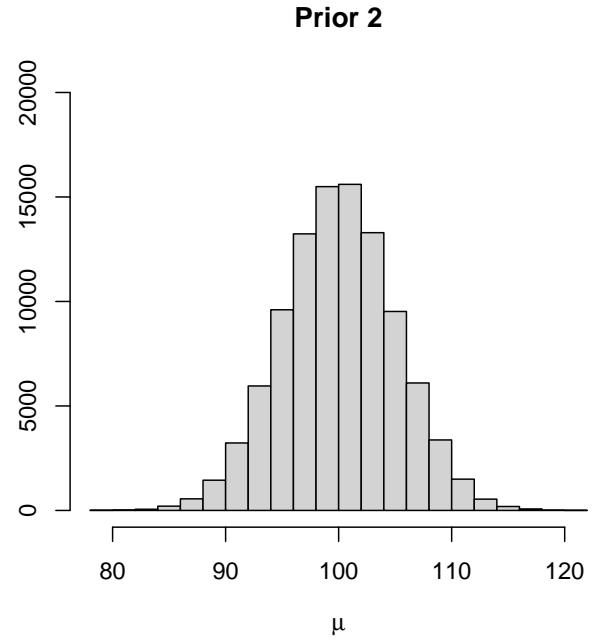
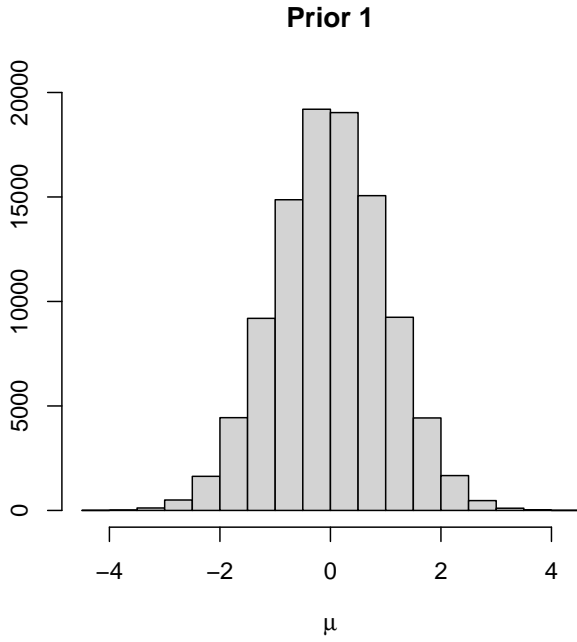
$$\begin{aligned}
 p(\mu|\mathbf{y}, 1) &\propto \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n\bar{y}^2 + \frac{1}{\sigma_0^2} \mu_0^2 - \frac{1}{\sigma_n^2} \mu_n^2 + \frac{1}{\sigma_n^2} (\mu^2 - 2\mu\mu_n + \mu_n^2)) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n\bar{y}^2 + \frac{1}{\sigma_0^2} \mu_0^2 - \frac{1}{\sigma_n^2} \mu_n^2 + \frac{1}{\sigma_n^2} (\mu - \mu_n)^2) \\
 &= \exp(-\frac{1}{2} (\sum_{i=1}^n (y_i - \bar{y})^2) + n\bar{y}^2 + \frac{1}{\sigma_0^2} \mu_0^2 - \frac{1}{\sigma_n^2} \mu_n^2) \times \exp(-\frac{1}{2\sigma_n^2} (\mu - \mu_n)^2) \\
 &\propto \exp(-\frac{1}{2\sigma_n^2} (\mu - \mu_n)^2)
 \end{aligned}$$

Hence we find that the posterior is proportional to the density of a distribution  $N(\mu_n, \sigma_n^2)$ . In the plot below, we further visualized the histograms of two prior distributions, where in Prior 1 we have  $\mu \sim N(0, 1)$  and for Prior 2 we specified  $\mu \sim N(100, 25)$ .

```

set.seed(123)
par(mfrow = c(1, 2))
hist(rnorm(1e+05, mean = 0, sd = 1), ylim = c(0, 20000), main = "Prior 1", xlab =
expression(mu),
     ylab = NA)
hist(rnorm(1e+05, mean = 100, sd = 5), ylim = c(0, 20000), main = "Prior 2", xlab =
expression(mu),
     ylab = NA)

```



## Estimating $\sigma^2$

**Q:** Suppose you have data  $\mathbf{y} \sim N(5, \sigma^2)$ , and want to estimate  $\sigma^2$ . Work with the precision,  $\sigma^{-2}$ , and specify a Gamma prior  $\sigma^{-2} \sim G(0.5, \eta)$  with single parameter  $\eta$ . Derive the posterior  $p(\sigma^2|\mathbf{y})$  by applying Bayes' theorem. Visualise the prior density for  $\eta \in \{0.01, 1, 100\}$ .

First, as we assume our data  $\mathbf{y} \sim N(5, \sigma^2)$ , the probability density of a single observation  $y_i$  is given by  $p(y_i|5, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp(-\frac{1}{2} \frac{(y_i-5)^2}{\sigma^2})$ . If we assume our data to be i.i.d., then the joint density (the likelihood function) is given by:

$$\begin{aligned} p(\mathbf{y}|5, \sigma^2) &= \prod_{i=1}^n p(y_i|5, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - 5)^2) \\ &\propto (\sigma^2)^{-\frac{n}{2}+1-1} \exp(-\frac{\sum_{i=1}^n (y_i - 5)^2/2}{\sigma^2}) \\ &= (\sigma^2)^{c-1} \exp(-\frac{d}{\sigma^2}) \end{aligned}$$

Hence, we find that the likelihood is proportional to an Inverse-Gamma distribution  $G^{-1}(c, d)$ , where  $c = -\frac{n}{2} + 1$  and  $d = \sum_{i=1}^n (y_i - 5)^2/2$ .

For the prior we assume that the precision follows a Gamma distribution, i.e.  $\sigma^{-2} \sim G(0.5, \eta)$ . Therefore, we have  $\sigma^2 \sim G^{-1}(0.5, \eta)$ .<sup>1</sup> Our prior for  $\sigma^2$  is then given by  $p(\sigma^2|0.5, \eta) = \frac{\sqrt{\eta}}{\Gamma(0.5)} (\sigma^2)^{-1.5} \exp(-\frac{\eta}{\sigma^2})$ .

To get to our posterior, we utilize Bayes' theorem, dropping all constant factors from our likelihood function and our prior in the process, since we are interested in proportionality:

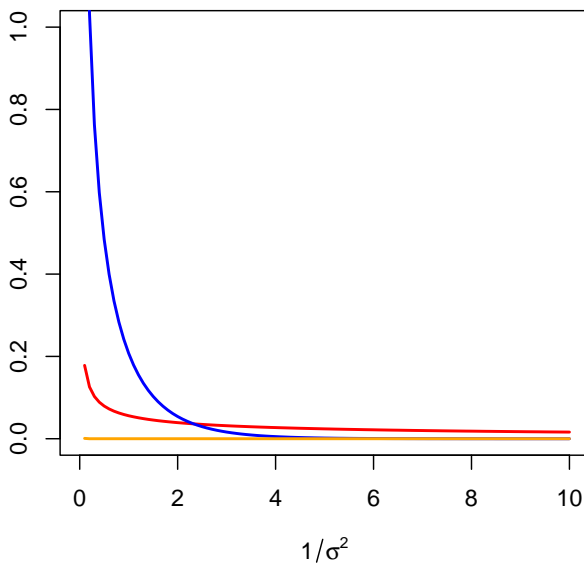
$$\begin{aligned} p(\sigma^2|\mathbf{y}, 5) &\propto p(\mathbf{y}|5, \sigma^2) \times p(\sigma^2|0.5, \eta) \\ &\propto (\sigma^2)^{-\frac{n}{2}} \exp(-\frac{\sum_{i=1}^n (y_i - 5)^2/2}{\sigma^2}) \times (\sigma^2)^{-1.5} \exp(-\frac{\eta}{\sigma^2}) \\ &= (\sigma^2)^{-(\frac{n}{2}+0.5)-1} \exp(-\frac{1}{\sigma^2} (\eta + \sum_{i=1}^n (y_i - 5)^2/2)) \\ &= (\sigma^2)^{-c_n-1} \exp(-\frac{d_n}{\sigma^2}) \end{aligned}$$

<sup>1</sup>See derivation of Inverse-Gamma from Gamma on [Wikipedia](#). Note that whereas for the Gamma distribution  $\beta$  is the rate parameter, it is the scale parameter for the Inverse Gamma distribution.

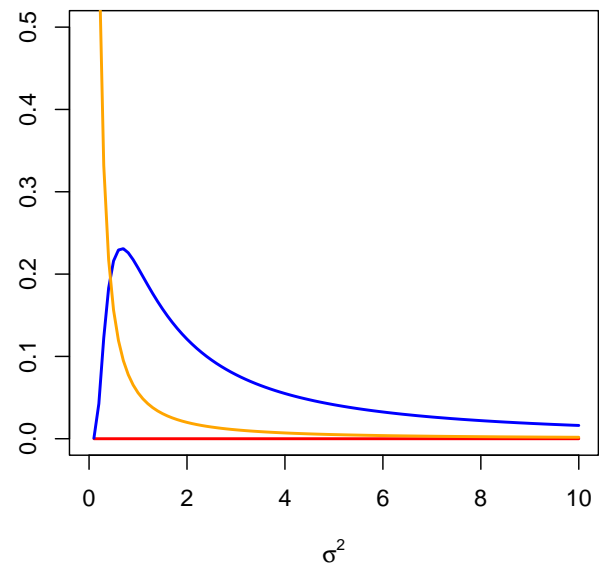
Therefore, the posterior is proportional to the pdf of an Inverse-Gamma distribution  $G^{-1}(c_n, d_n)$ , where  $c_n = \frac{n}{2} + 0.5$  and  $d_n = \eta + \sum_{i=1}^n (y_i - 5)^2 / 2$ . In the plot below, we visualised the prior density for  $\eta \in \{0.01, 1, 100\}$  for both the precision  $\sigma^{-2}$  (on the left) as well as the variance  $\sigma^2$  (on the right).

```
cols <- c("red", "blue", "orange")
eta <- c(0.01, 1, 100)
par(mfrow = c(1, 2))
plot(NA, NA, xlim = c(0, 10), ylim = c(0, 1), main = "Prior: Precision", xlab =
  expression(1/sigma^2),
  ylab = NA)
for (i in 1:length(eta)) {
  curve(dgamma(x, shape = 0.5, rate = eta[i]), from = 0, to = 10, col = cols[i],
    add = T, lwd = 2)
}
plot(NA, NA, xlim = c(0, 10), ylim = c(0, 0.5), main = "Prior: Variance", xlab =
  expression(sigma^2),
  ylab = NA)
for (i in 1:length(eta)) {
  curve(invgamma::dinvgamma(x, shape = 0.5, scale = eta[i]), from = 0, to = 10,
    col = cols[i], add = T, lwd = 2)
}
legend(3, 0.9, c("eta1=0.01", "eta2=1", "eta3=100"), lwd = 2, col = cols)
```

**Prior: Precision**



**Prior: Variance**



## Including $\eta$

First, the hyperprior should be chosen such that the support of this distribution fits the values that  $\eta$  is allowed to take. As the scale parameter of an Inverse-Gamma distribution has to be larger than 0, we need to find a hyperprior such that random draws result in  $\eta \in (0, \infty)$ . Several distributions such as the Gamma, Inverse-Gamma, or the Log-Normal distribution fulfill this criteria. Now from the type of data that we are working with, we may have some kind of an idea as to where  $\sigma^2$  should be. Let's say that we are quite sure that  $\sigma^2$  is most likely to be around 1, but we also want to allow for larger error terms. Then this would correspond to the density of an Inverse-Gamma function with  $\alpha = 0.5$  (which is fixed by assumption) and a scale parameter  $\eta$  that is most of the time somewhere between 0.5 and 1. This can be achieved by drawing from  $\eta \sim \text{Lognormal}(-0.4, 0.25)$  (see the density for this distribution on the left hand side of the plot below). The prior  $\sigma^2 | \eta$  is visualized on the right hand side of the plot below.

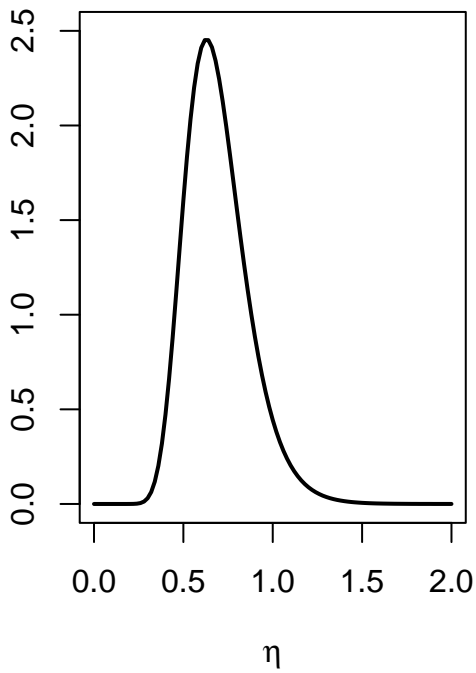
```
set.seed(123)
eta.prior <- data.frame(eta = rlnorm(1000, meanlog = -0.4, sdlog = 0.25), sigma2 = NA)
for (i in 1:nrow(eta.prior)) {
  eta.prior$sigma2[i] <- invgamma::rinvgamma(1, shape = 0.5, scale = eta.prior$eta[i])
}
```

```

par(mfrow = c(1, 2))
plot(NA, NA, xlim = c(0, 2), ylim = c(0, 2.5), main = "Density: Lognormal(-0.4, 0.25)",
     xlab = expression(eta), ylab = NA)
curve(dlnorm(x, meanlog = -0.4, sdlog = 0.25), from = 0, to = 2, add = T, lwd = 2)
plot(density(eta.prior$sigma2), main = expression(paste("Prior: ", sigma^2, "|",
eta)), xlab = expression(sigma^2), ylab = NA)

```

**Density: Lognormal(-0.4, 0.25)**



**Prior:  $\sigma^2|\eta$**

