Student: 670009416

Candidate: 050662

MTH3028

Statistical Inference

CW2: Likelihood & Computational Inference.

"Age of the Universe"

Contents:

```
require(tidyverse); set.seed(670009416)
```

Introduction: The file `supernovae_mth3028.csv` contains velocity (Mpc/Ga) and distance (Mpc) data about 36 Type 1A supernovae. Let y_i and x_i denote the velocity and distance of a supernova, respectively. The velocity is to be modelled as: $Y_i = \beta x_i + \epsilon_i$, with density function:

$$f(y_i; \beta, \tau) = (2\pi e^\tau)^{-\frac{1}{2}} e^{-\frac{1}{2} e^{-\tau}(y_i - \beta x_i)^2}$$

where β and τ are parameters. Let $\gamma = 1/\beta$.

Question 1: Prelims

a) We begin by visualizing the the speed and distance of the 36 supernovae. We also conduct some data wrangling to suit the needs of the study; transforming the string list to numeric data.

```
supern <- read.csv(file="supernovae_mth3028.csv", header=FALSE)[2:37,] # Only Relevant Data
super <- data.frame(matrix(as.numeric(unlist(supern)), nrow=36, ncol=2)) # Numeric data frame
colnames(super) <- c("velocity", "distance")
plot(super$distance, super$velocity,
     xlab="Distance (Mpc)",ylab="Velocity (Mpc/Ga)",
     xlim=c(0, max(super$distance)*1.05), ylim=c(0,max(super$velocity)*1.05), pch=3)
grid(lwd = 0.75); axis(1)
```
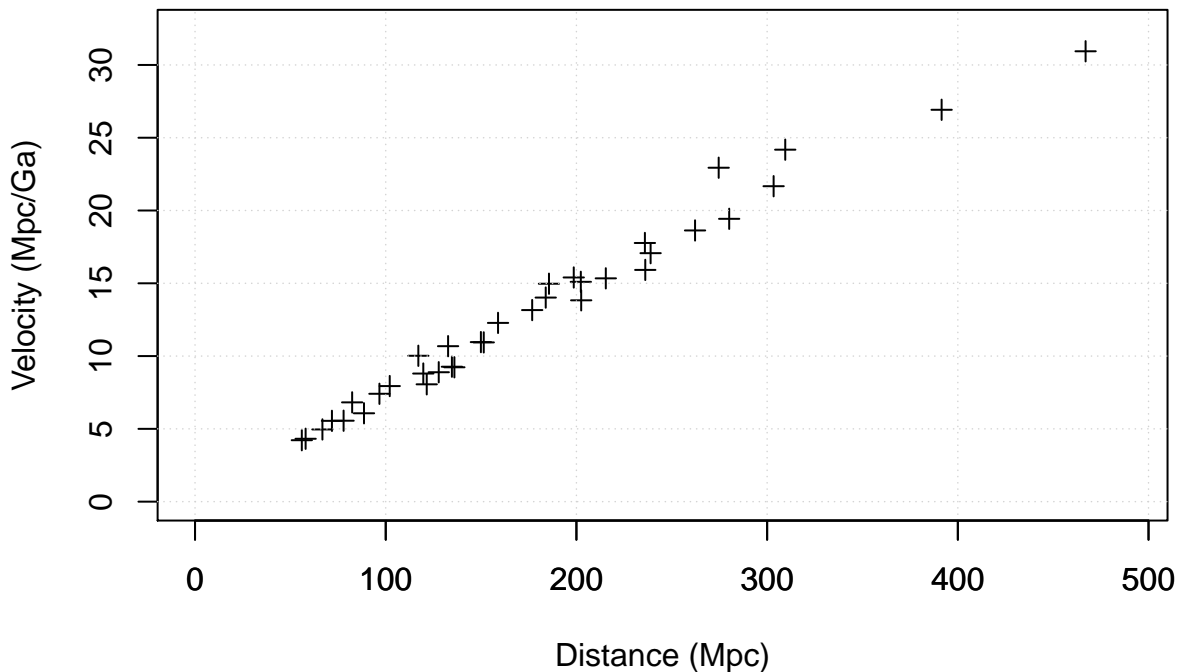

Figure 1: 36 Type 1A Supernovae Data

b) The assumption here, that $y_i = \beta x_i$, results that the velocity of a star undergoing a supernova is linearly related to the distance of that star from our planet, which is equal to $\frac{y_i}{x_i} = \beta$. This fraction is the same for all stars, and means β is equal to velocity over distance. This equal to inverse time; hence our $\gamma = \frac{1}{\beta}$ leads to a measure of time, which is the same for each star. Thus, the observations suggests that all stars ave moved from an original point at the same initial time, γ years ago. Hence, this is the time since the "big bang" moment, and is the "age" of the universe.

1

Question 2: Likelihood Function

Here we derive the likelihood function, its log form and its relevant derivatives.

a) For the likelihood function, $L(\beta, \tau; \mathbf{Y})$, we suppose all Y_i are independent, with density function $f(y_i; \beta, \tau)$:

$$
\begin{aligned}
L(\beta, \tau; \mathbf{Y}) &= \prod_{i=1}^{n} (2\pi e^{\tau})^{-\frac{1}{2}} e^{-\frac{1}{2} e^{-\tau} (Y_i - \beta x_i)^2} \\
&= (2\pi e^{\tau})^{-\frac{n}{2}} \exp\left\{ -\frac{e^{-\tau}}{2} \sum_{i=1}^{n} (Y_i - \beta x_i)^2 \right\}
\end{aligned}
$$

b) Log-Likelihood, $l(\beta, \tau) = \log\left[L(\beta, \tau; \mathbf{Y})\right]$

$$
l(\beta, \tau) = -\frac{n}{2} \left[\log(2\pi) + \tau\right] - \frac{e^{-\tau}}{2} \sum_{i=1}^{n} (Y_i - \beta x_i)^2
$$

c) Score function $U(\beta, \tau) = \left(\frac{\partial l(\beta, \tau)}{\partial \beta}, \frac{\partial l(\beta, \tau)}{\partial \tau} \right)$, where

$$
\begin{aligned}
\frac{\partial}{\partial \beta} l(\beta, \tau) &= e^{-\tau} \sum_{i=1}^{n} x_i (Y_i - \beta x_i) \\
\frac{\partial}{\partial \tau} l(\beta, \tau) &= -\frac{n}{2} + \frac{e^{-\tau}}{2} \sum_{i=1}^{n} (Y_i - \beta x_i)^2
\end{aligned}
$$

d) Observed Information $J(\beta, \tau)$, where

$$
\begin{aligned}
J(\beta, \tau) &= \begin{bmatrix} -\frac{\partial^2 l}{\partial \beta^2} & -\frac{\partial^2 l}{\partial \beta \partial \tau} \\ -\frac{\partial^2 l}{\partial \beta \partial \tau} & -\frac{\partial^2 l}{\partial \tau^2} \end{bmatrix} = \begin{bmatrix} e^{-\tau} \sum_{i=1}^{n} x_i^2 & e^{-\tau} \sum_{i=1}^{n} x_i (Y_i - \beta x_i) \\ e^{-\tau} \sum_{i=1}^{n} x_i (Y_i - \beta x_i) & \frac{e^{-\tau}}{2} \sum_{i=1}^{n} (Y_i - \beta x_i)^2 \end{bmatrix} \\
\frac{\partial^2 l}{\partial \beta^2} &= -e^{-\tau} \sum_{i=1}^{n} x_i^2 \\
\frac{\partial^2 l}{\partial \beta \partial \tau} &= -e^{-\tau} \sum_{i=1}^{n} x_i (Y_i - \beta x_i) \\
\frac{\partial^2 l}{\partial \tau^2} &= -\frac{e^{-\tau}}{2} \sum_{i=1}^{n} (Y_i - \beta x_i)^2
\end{aligned}
$$

e) Expected Information $I(\beta, \tau) = -E\left[J(\beta, \tau)\right]$. Importantly, we have that $E(Y_i) = \beta x_i$, and $E(Y_i - \beta x_i)^2 = e^{\tau}$, as e^{τ} is the variance, and $E(Y - E(Y))^2 = Var(Y)$, by definition. This leads to

$$
\begin{aligned}
I(\beta, \tau) &= \begin{bmatrix} -E\left[\frac{\partial^2 l}{\partial \beta^2}\right] & -E\left[\frac{\partial^2 l}{\partial \beta \partial \tau}\right] \\ -E\left[\frac{\partial^2 l}{\partial \beta \partial \tau}\right] & -E\left[\frac{\partial^2 l}{\partial \tau^2}\right] \end{bmatrix} \\
&= \begin{bmatrix} e^{-\tau} \sum_{i=1}^{n} x_i^2 & 0 \\ 0 & \frac{n}{2} \end{bmatrix}
\end{aligned}
$$

Where, in particular,

$$
-E\left[\frac{\partial^2 l}{\partial \tau^2}\right] = \frac{1}{2e^{\tau}} \sum_{i=1}^{n} E(Y_i - \beta x_i)^2 = \frac{1}{2e^{\tau}} n e^{\tau} = \frac{n}{2}
$$

Question 3: Estimation of β and τ

a) We derive the MLE estimates for the parameters by setting $U(\hat{\beta}, \hat{\tau}; \mathbf{Y}) = \mathbf{0}$:

$$e^{-\hat{\tau}} \sum_i x_i(Y_i - \hat{\beta}x_i) = 0 \quad \rightarrow \quad \sum_i [x_iY_i] - \hat{\beta}\sum_i [x_i^2] = 0 \quad \rightarrow \quad \hat{\beta} = \frac{\sum_i [x_iY_i]}{\sum_i [x_i^2]}$$

$$-\frac{n}{2} + \frac{e^{-\hat{\tau}}}{2}\sum_i(Y_i - \beta x_i)^2 = 0 \quad \rightarrow \quad e^{-\hat{\tau}}\sum_i(Y_i - \beta x_i)^2 = n \quad \rightarrow \quad \hat{\tau} = \log\left[\frac{1}{n}\sum_i(Y_i - \hat{\beta}x_i)^2\right]$$

Importantly, we see if we substitute these estimates into he observed second derivative's elements $(l''(\beta, \tau; \mathbf{y}))$, then they are all negative. In particular, $\hat{\tau}$ is reduced to the argument of its logarithm since it's a power of the exponential, which is clearly positive, so the overall terms are negative. Since $e^{-\tau}$ is a coefficient in all these term, the aforementioned is true for all second derivative elements. Hence, these estimates are indeed the maximum values possible.

Now, by putting these into R, we can compute the MLE for β and τ from the data:

```
n = nrow(super) # n=36 For All Supernovae
x = super$distance
y = super$velocity
beta_hat = sum(x*y)/sum(x^2)
tau_hat = log(mean((y-beta_hat*x)^2))
# Final MLE for all parameters:
print(c("Beta-Hat"=beta_hat, "Tau-Hat"=tau_hat))
```

```
##   Beta-Hat     Tau-Hat
## 0.07227653 0.04125402
```

b) Asymptotic distribution of the MLE is

$$\begin{pmatrix} \hat{\beta} \\ \hat{\tau} \end{pmatrix} \dot{\sim} N \left[\begin{pmatrix} \beta \\ \tau \end{pmatrix}, \begin{pmatrix} \frac{e^{\tau}}{\sum_{i=1}^n x_i^2} & 0 \\ 0 & \frac{2}{n} \end{pmatrix} \right]$$

Hence, we have that $Var(\hat{\beta}) = e^{\tau}\left[\sum_i x_i^2\right]^{-1}$, and $Var(\hat{\tau}) = \frac{2}{n}$. Hence, for $se(\hat{\tau})$ and $se(\hat{\beta})$ we can compute:

```
se_tau_hat = sqrt(2/n)
se_beta_hat = sqrt({exp(1)^{tau_hat}}/ (sum(x^2)))
print(c("se(Beta-Hat)" = se_beta_hat, "se(Tau-Hat)" = se_tau_hat))
```

```
## se(Beta-Hat)   se(Tau-Hat)
## 0.0008453752 0.2357022604
```

Question 4: Age of the Universe

a) The point estimate for γ is $\hat{\gamma} = \frac{1}{\hat{\beta}}$. This can be computed as:

```
gamma_hat = 1/beta_hat
print(c("Gamma_Hat"=gamma_hat))
```

```
## Gamma_Hat
##  13.83575
```

b) By the asymptotic normal distribution of the MLE we have that $E(\hat{\beta}) = \beta$ and $var(\hat{\beta}) \approx \frac{e^{\tau}}{\sum_i x_i^2}$.

Now, by using $h(\beta) = 1/\beta$, we can use the delta method to approximate the variance of $\hat{\gamma}$:

$$var(\hat{\gamma}) \approx var(\hat{\beta}) \left[h'(\beta)\right]^2 = \frac{e^{\tau}}{\sum_i x_i^2} \frac{1}{\beta^4}$$

```
var_gamma_hat = ((exp(1)^tau_hat)/sum(x^2)) * (1/beta_hat^4)
print(c("var(Gamma-Hat)" = var_gamma_hat))
```

```
## var(Gamma-Hat)
##      0.02618845
```

c) We can print out estimate of the age of the universe $\hat{\gamma}$ and print its associated standard error:

```
gamma_hat_se = sqrt(var_gamma_hat)
print(c("Age of the Universe"=gamma_hat, "Sd.Error"=gamma_hat_se))
```

```
## Age of the Universe          Sd.Error
##          13.8357492         0.1618285
```

Hence the universe is estimated to be 13.8357 ± 0.1618 billions years old ($4d.p.$).

We can also develop a bias-corrected estimate ($\tilde{\gamma}$) for the age of the universe:

$$E(\hat{\gamma}) - \gamma \approx \frac{1}{2} var(\hat{\beta}) h''(\beta) = \frac{e^{\tau}}{\sum_i x_i^2} \frac{1}{\beta^3} \qquad \rightarrow \qquad \tilde{\gamma} = \hat{\gamma} - \frac{e^{\tau}}{\hat{\beta}^3 \sum_i x_i^2} = \hat{\gamma} \left(1 - \hat{\gamma}^2 \frac{e^{\tau}}{\sum_i x_i^2}\right)$$

```
gamma_tilde = gamma_hat - (exp(1)^tau_hat)/((beta_hat^3)*sum(x^2))
gamma2 = gamma_hat * (1 - gamma_hat^2 * exp(1)^tau_hat / sum(x^2))
print(c("Age of the Universe (Bias Corrected)"=gamma_tilde))
```

```
## Age of the Universe (Bias Corrected)
##                             13.83386
```

Hence the universe is, biased corrected, estimated be 13.833 billions years old. Also, the bias correction will have little effect on he variance; hence $\tilde{\gamma}$ will have similar variance to $\hat{\gamma}$.

Question 5: Testing the Universe Age Estimate

Now let $\theta = (\beta, \tau)$, the general vectorised form of the parameters.

a) We have the test H_0: $\theta = \left(\frac{1}{13.5}, 0\right)$, against H_1: $\theta \neq \left(\frac{1}{13.5}, 0\right)$, with size 0.05. The Score test statistic is $S = U(\theta_0)^T I(\theta_0)^{-1} U(\theta_0)$. Hence we can undertake this test in R:

```
alpha = 0.05
beta_0 = 1/(13.5)
tau_0 = 0
theta_0 = c(beta_0, tau_0)
U_thetaz = c({exp(-tau_0)}*sum(x*(y-beta_0*x)), -{n/2}+{exp(-tau_0)/2}*sum((y-beta_0*x)^2))
I_hat_inv = diag(c(exp(1)^(tau_0)/sum(x^2), 2/n))
S = t(U_thetaz) %*% I_hat_inv %*% U_thetaz
print(c("Score Test Statistic (S)"= S))
```

```
## Score Test Statistic (S)
##                 5.250365
```

Now we can calculate the critical region for our test, by calculating the $1 - \alpha$-quantile of the χ_2^2 distribution:

```
lb2 = qchisq(1-alpha,df=2)
paste0("Critical Value: ", lb2)
```

```
## [1] "Critical Value: 5.99146454710798"
```

```
paste0("Critical Region =  (",round(lb2, 3), ", inf), to 3d.p.")
```

```
## [1] "Critical Region =  (5.991, inf), to 3d.p."
```

Hence, our S is clearly not in the critical region, so the score test does not reject the null hypothesis at the 0.05 size.

b) For the profile-likelihood $l_p(\beta) = l(\beta, \hat{\tau}(\beta))$, we must first maximise the value of τ:

$$\frac{\partial l}{\partial \tau} = -\frac{n}{2} + \frac{1}{2e^\tau} \sum_{i=1}^{n} (Y_i - \beta x_i)^2 = 0$$

$$\therefore \quad \hat{\tau}(\beta) = \log\left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - \beta x_i)^2\right]$$

$$\rightarrow \quad l_{p_\beta}(\beta) = -\frac{n}{2} - \frac{n}{2} \log\left[\frac{2\pi}{n} \sum_{i=1}^{n} (Y_i - \beta x_i)^2\right]$$

Now, to develop the likelihood ratio Λ:

$$\Lambda_{p_\beta} = \frac{\sup_{(\beta,\tau)\in\Theta_0} L[(\beta,\tau); \mathbf{y}]}{\sup_{(\beta,\tau)} L[(\beta,\tau); \mathbf{y}]} = \frac{L[(\beta_0, \hat{\tau}_0); \mathbf{y}]}{L[(\hat{\beta}, \hat{\tau}); \mathbf{y}]}$$

Where, $\hat{\tau}_0 = \log\left[\frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 x_i)^2\right]$.
Now, the test statistic of the β-profile likelihood ratio test statistic is

$$-2 \log \Lambda_{p_\beta} = -2 \left(l_{p_\beta}[(\beta_0, \hat{\tau}_o); \mathbf{y}] - l_{p_\beta}[(\hat{\beta}, \hat{\tau}); \mathbf{y}]\right)$$

Importantly, if we consider the $\hat{\tau}(\beta_0) = \hat{\tau}_0$ from above, then the LRT statistic becomes:

$$-2 \log \Lambda_{p_\beta} = n \left[\hat{\tau}_0 - \hat{\tau}\right]$$

Now, we can develop this test in R:

```r
alpha = 0.05
beta_0 = 1/(13.5)
tau_0_hat = log(mean((y-beta_0*x)^2))

like_p_beta_0 = -n/2-(n/2)*log(2*pi*mean( (y-beta_0*x)^2))
like_p_beta_hat = -n/2-(n/2)*log(2*pi*mean( (y-beta_hat*x)^2))
LRT_stat = -2*(like_p_beta_0 - like_p_beta_hat)

lb2 = qchisq(1-alpha,df=1) # One-Tailed Chi squared
paste0("Critical Region = (",round(lb2, 3), ", inf), to 3d.p.")
```

```
## [1] "Critical Region = (3.841, inf), to 3d.p."
```

```r
paste0("LRT Statistic = ", round(LRT_stat, 3))
```

```
## [1] "LRT Statistic = 4.259"
```

Clearly, our LRT Statistic is in the critical region, when comparing to the $1 - \alpha$ quantile of the χ_1^2 distribution. Hence, we reject H_0 in favour of the alternative hypothesis, at size 0.05.

c) From the critical region, we have that

$$-2\log\Lambda = n(\hat{\tau}_0 - \hat{\tau}) \geq c$$

where c denotes the 95-percentile of the χ_1^2 distribution.

Now, for the confidence interval, $S(\mathbf{Y})$, we want to include all the values of β_0 that are not included in the critical region. Hence we have that

$$S(\mathbf{Y}) \quad = \quad \left\{ \beta : \log\left(\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta x_i)^2 \right) \geq \frac{c}{n} + \hat{\tau} \right\}$$

Now, using that $y_i = \hat{\beta}x_i$,

$$
\begin{aligned}
ne^{\frac{c}{n}+\hat{\tau}} \quad &\leq \quad \sum_{i=1}^{n}y_i^2 + \beta^2\sum_{i=1}^{n}x_i^2 - \beta\sum_{i=1}^{n}y_ix_i \\
&\leq \quad \sum_{i=1}^{n}y_i^2 + \beta^2\sum_{i=1}^{n}x_i^2 - \beta\hat{\beta}\sum_{i=1}^{n}x_i^2 + \left[\hat{\beta}^2\sum_{i=1}^{n}x_i^2 - \hat{\beta}^2\sum_{i=1}^{n}x_i^2\right] \\
&\leq \quad \sum_{i=1}^{n}y_i^2 + (\beta - \hat{\beta})^2\sum_{i=1}^{n}x_i^2 - \hat{\beta}^2\sum_{i=1}^{n}x_i^2 \\
\therefore \quad (\beta - \hat{\beta})^2 \quad &\geq \quad \hat{\beta}^2 + \frac{ne^{\frac{c}{n}+\hat{\tau}} - \sum_{i=1}^{n}y_i^2}{\sum_{i=1}^{n}x_i^2}
\end{aligned}
$$

This leads to the 95% confidence interval for β:

$$\left(\hat{\beta} - \sqrt{\hat{\beta}^2 + \frac{ne^{\frac{c}{n}+\hat{\tau}} - \sum_{i=1}^{n}y_i^2}{\sum_{i=1}^{n}x_i^2}}, \quad \hat{\beta} + \sqrt{\hat{\beta}^2 + \frac{ne^{\frac{c}{n}+\hat{\tau}} - \sum_{i=1}^{n}y_i^2}{\sum_{i=1}^{n}x_i^2}} \right)$$

```
ch = qchisq(1-alpha, 1)
inside = beta_hat^2 + (n*(exp(ch/n+tau_hat))-sum(y^2))/sum(x^2)
paste0("(", beta_hat - sqrt(inside), ", ", beta_hat + sqrt(inside), ")")
```

```
## [1] "(0.0705744296073721, 0.0739786385732296)"
```

```
paste0("Confidence interval for Beta: (",
       round(beta_hat - sqrt(inside), 5), ", ",
       round(beta_hat + sqrt(inside), 5), "), rounded.")
```

```
## [1] "Confidence interval for Beta: (0.07057, 0.07398), rounded."
```

c) Now, in order to calculate the the coverage of the previous confidence intervals, we use Monte Carlos simulation. We repeatedly simulate new velocity values (y) and then estimate the confidence interval:

```
beta_0 = 1/(13.5)
tau_0 = 1
ch = qchisq(1-alpha, 1)

monte_carlo = replicate(1e4, {
y_new = beta_0*x + rnorm(n, 0, sqrt(exp(tau_0)))
beta_new = sum(x*y_new)/sum(x^2)
tau_new = log(mean((y_new-beta_new*x)^2))
argument = beta_new^2 + (n*(exp(ch/n+tau_new))-sum(y_new^2))/sum(x^2)
c(beta_new - sqrt(argument), beta_new + sqrt(argument))
})
```

Now, we simply count how many times the confidence intervals don't cover the value of β_0.

```
mean(beta_0>monte_carlo[1,] & beta_0<monte_carlo[2,])
```

```
## [1] 0.9436
```

Hence, we see that our coverage is slightly smaller than the desired coverage of 0.95, but its is close to his desired value.

Question 6: Universe Age Estimate's Robustenss

a) Bootstrap Versions of γ

(i) Parametric Bootstrap for $\hat{\gamma}$, 10^5 repetitions. We develop this estimate by taking 36 samples $y_i^\star = \hat{\beta}x_i + \epsilon_i$, with $\epsilon \sim N(0, e^{\hat{\tau}})$ are independent and identically distributed. We then calculate $\hat{\gamma}_i^\star$ estimates from each of these $y_i^\star$ and the x_i data.

```
set.seed(670009416+1)
gamma_hat_parametric = replicate(1e5, {
  y_star = rnorm(n, beta_hat * x, sqrt(exp(1)^tau_hat))
  beta_star = sum(y_star * x)/sum(x^2)
  tau_star = log(((y_star-beta_star*x)^2)/n)
  1/beta_star})
```

(ii) Semi-Parametric (error-sampling) Bootstrap for $\hat{\gamma}$, 10^5 repetitions. We develop an algorithm to calculate residuals $res_i = y_i - \hat{\beta}x_i$. Then we drawn n samples of $res_i^\star$ (potentially non-unique) and calculate a new $y_i^\star = \hat{\beta}x_i + res_i$. When we can estimate $\hat{\gamma}^\star$ from the calculated $y_i^\star$ and x_i.

```
set.seed(670009416+2)
res = y-beta_hat*x   # Forming Residuals
gamma_hat_semiparametric = replicate(1e5, {
  res_star = sample(res, rep=TRUE)   # Sampling Residuals
  y_star = beta_hat*x + res_star # Error Sampled
  beta_star = sum(x*y_star)/sum(x^2)
  1/beta_star}) # Gamma hat
```

(iii) Non-Parametric Bootstrap for $\hat{\gamma}$, with 10^5 repetitions. For this we develop an algorithm that draws n (potentially non-unique) pairs of x and y from the data set of 36 stars and then estimates a $\hat{\gamma}^\star$, from the calculated $y^\star$

```
set.seed(670009416+3)
gamma_hat_nonparametric = replicate(1e5,{
  i_star = sample(1:n,rep=TRUE)
  x_star = x[i_star]
  y_star = y[i_star]
  beta_star = sum(x_star*y_star)/sum(x_star^2)
  1 / beta_star
})
```

Now to derive the bootstrap 95% confidence interval for γ from our three bootstrap simulations:

```
gamma_hat_bootstrap = cbind("Parametric" = gamma_hat_parametric,
                            "SemiParametric" = gamma_hat_semiparametric,
                            "NonParametric" = gamma_hat_nonparametric)
apply(gamma_hat_bootstrap, 2,
      function(g) 2*gamma_hat-quantile(g, c(0.975, 0.025))) # Variable Quantile
```

```
##        Parametric SemiParametric NonParametric
## 97.5%   13.51166       13.65153       13.40844
## 2.5%    14.14921       14.27023       14.34688
```

8

b) Now we can visualise the density plots of all of these bootstrap estimates:

```
density_parametric = density(gamma_hat_parametric)
density_nonparametric = density(gamma_hat_nonparametric)
density_semiparametric = density(gamma_hat_semiparametric)
with(density_parametric, plot(x, y, type='l', xlab=expression(hat(gamma)), ylab='Density Values',
                    xlim=c(min(density_parametric$x), max(density_parametric$x)), lty=1,
                    ylim=c(min(density_parametric$y), max(density_parametric$y)*1.05)))
with(density_semiparametric, lines(x, y, lty=2))
with(density_nonparametric, lines(x, y, lty=3))
legend('topright', c('Parametric', 'Semi-Parametric', 'Non-Parametric'), lty=1:3, bty='n')
grid(lwd = 0.75); axis(1)
```
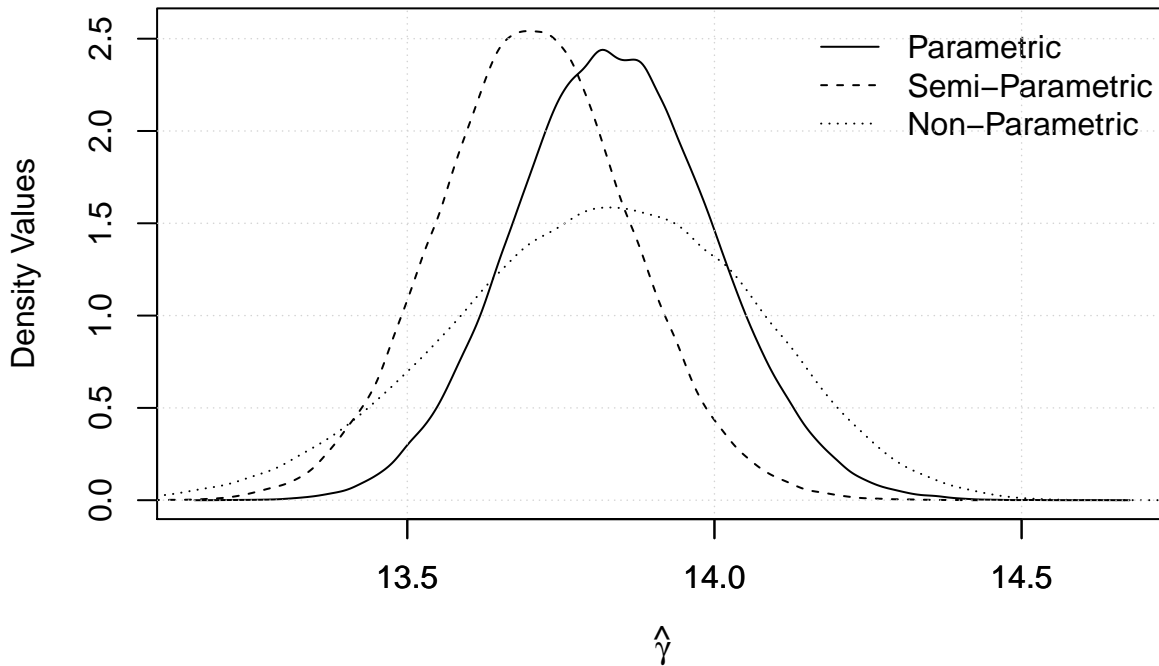

Figure 2: Gamma Bootstrap Estimates

All densities are approximately centred on the same value between 13.5 and 14 billion years (Ga).
The non-parametric bootstrap distribution has greater variance than the other two, because the distribution is wider, with a lower peak; the Parametric and semi-parametric bootstrap distributions have similar variance, with the parametric estimate having slightly more apparent variance.
Also, its clear that the semi-parametric distribution has a lower mean than the other two, leading to a lower expected age of the universe; the semi-parametric and nonparametric estimates have an apparently higher mean, though its difficult to distinguish which is higher form the plot.

c) The given best estimate for the age of the universe, 13.799 Billion years (Ga, ± 0.021), lies at an approximately medial point in our bootstrap distributions in part b, as well as the confidence intervals in part a. Hence this estimate agrees with these calculations. Also, this value implies a value of β that lies in the confidence set estimates in question 5. Importantly, the standard error given in this estimate, of 0.021 billion years, is smaller than the standard errors of the bootstrap distributions. The likely reason for this better estimate is that it was developed with more data than we have used in this study. Also, the estimate may have been developed with a different density function than may capture the variability of the data better and lead to a better estimate.