Analysis of Environmental Data LAB 4

Olivia Dinkelacker

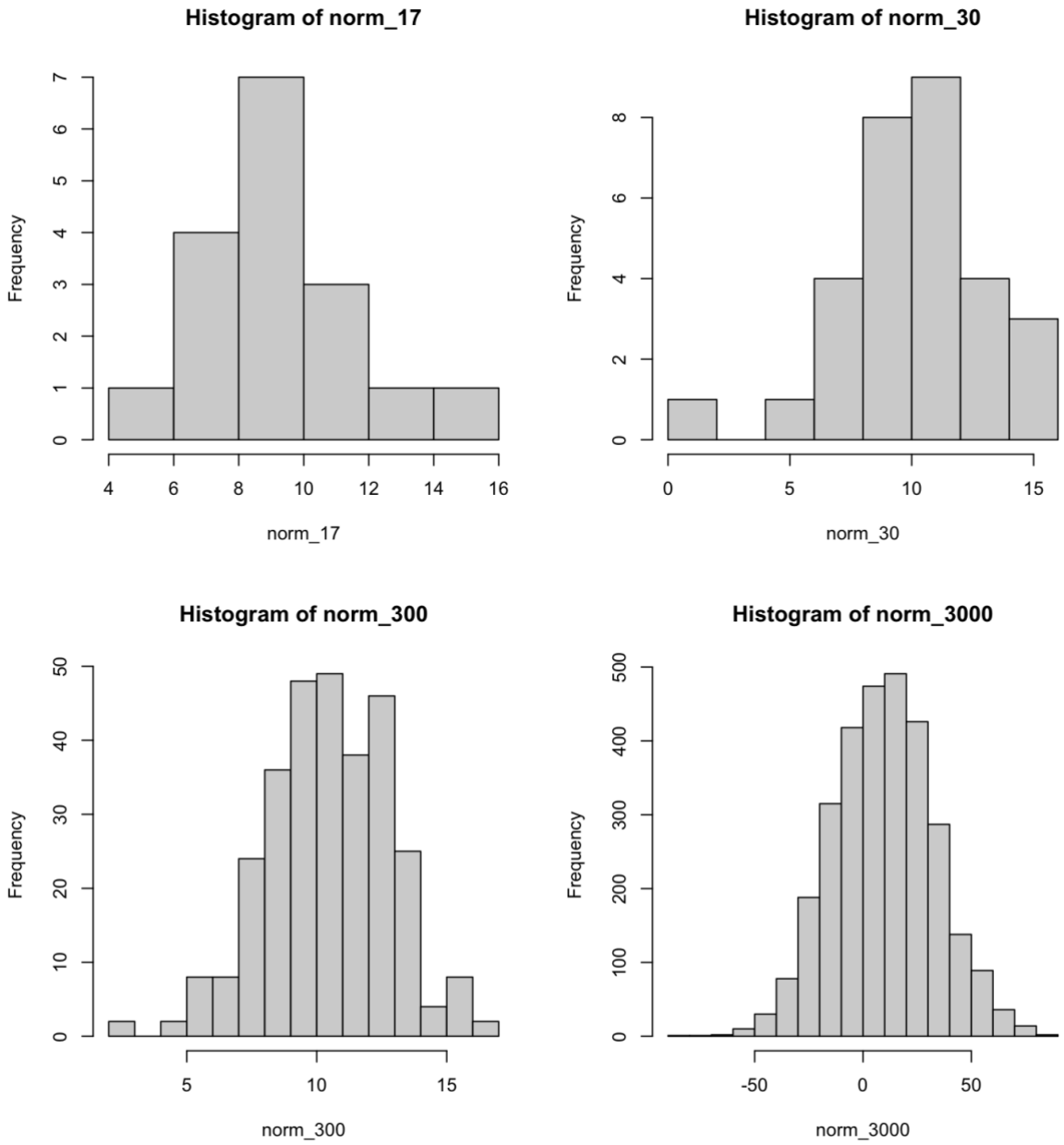
Q1 Vector Code

```
norm_17 = rnorm(n = 17, mean = 10.4, sd =2.4)

norm_30 = rnorm(n = 30, mean = 10.4, sd = 2.4)

norm_300 = rnorm(n = 300, mean = 10.4, sd = 2.4)

norm_3000 = rnorm(n = 3000, mean = 10.4, sd = 24)
```

Q2 Histogram Code

```
par(mfrow= c(2,2))

hist(norm_17)

hist(norm_30)

hist(norm_300)

hist(norm_3000)

    #saving as png

    require(here)

    png(

      filename = here("lab_04_hist_01.png"),

      width = 1500, height = 1600,

      res = 180, units = "px")

    dev.off() #save file
```

Q3 Histogram Figure



Q4 Histogram Shapes 1

- (2 pts.) Qualitatively describe the differences among the histograms.

The two upper plots are skewed in opposite ways. The first plot on the bottom right is slightly skewed, whereas the last plot seems to be fairly normally distributed (bump-shaped).

Q5 Histogram Shapes 2

- Explain *why* the shapes of the histograms are different.

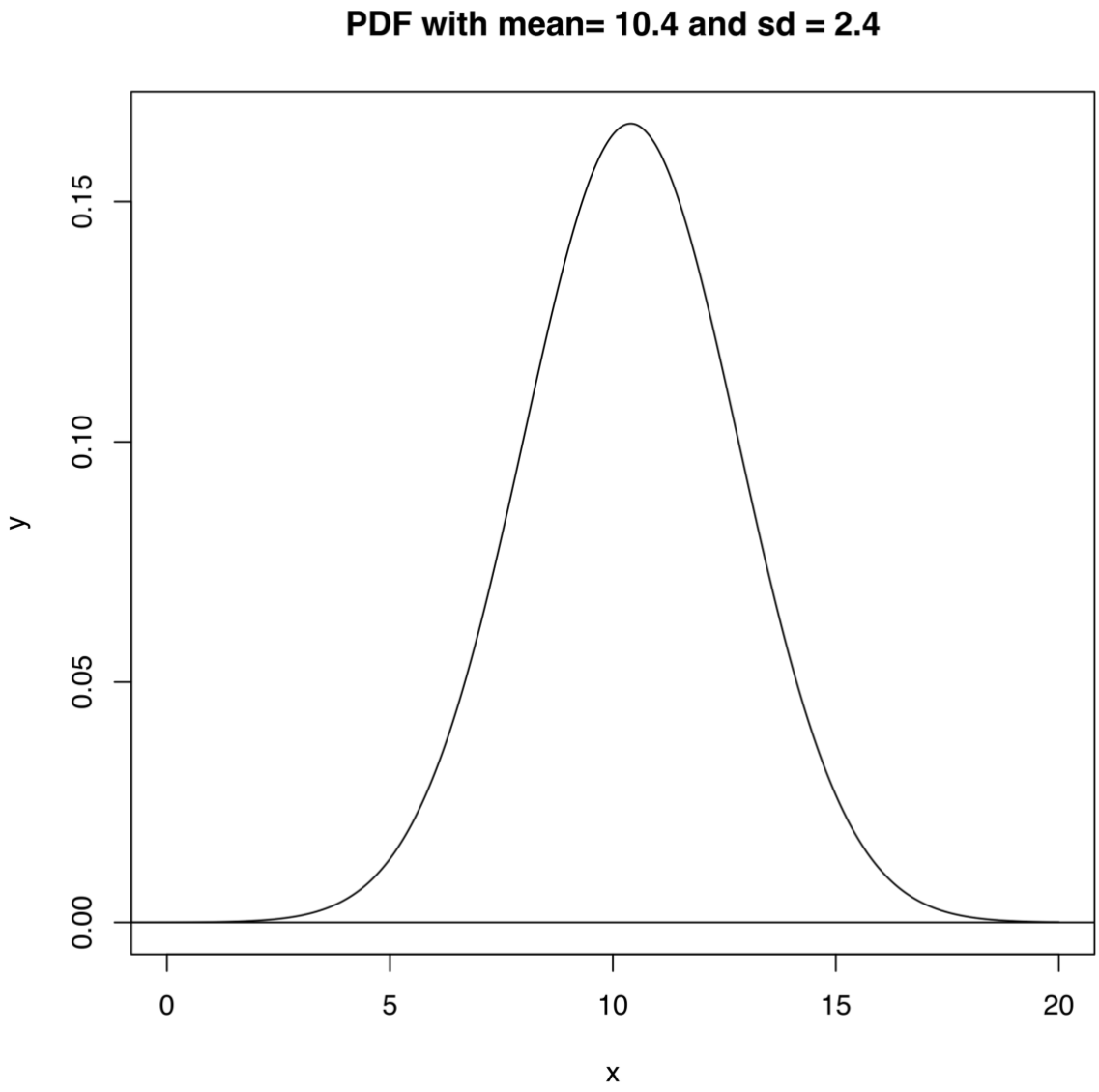
The sample size n = 17 and n = 30 is very small. Small samples are often ubject to sampling error and are non-representative samples. If we have more randomly generated numbers, the histograms look more uniform. The bigger the sample size the more normally distributed is the data, because the less samples the higher the uncertainty.

Q6: What are the parameters and their values for the standard Normal distribution?
The mean = 0 and standard deviation = 1.

Q7 Density Figure Code

```
x = seq(-20, 20, length.out = 1000)
y = dnorm(x, mean = 10.4, sd = 2.4)


plot(x, y, main = "PDF with mean= 10.4 and sd = 2.4", type = "l", xlim = c(0, 20))
abline(h = 0)
```

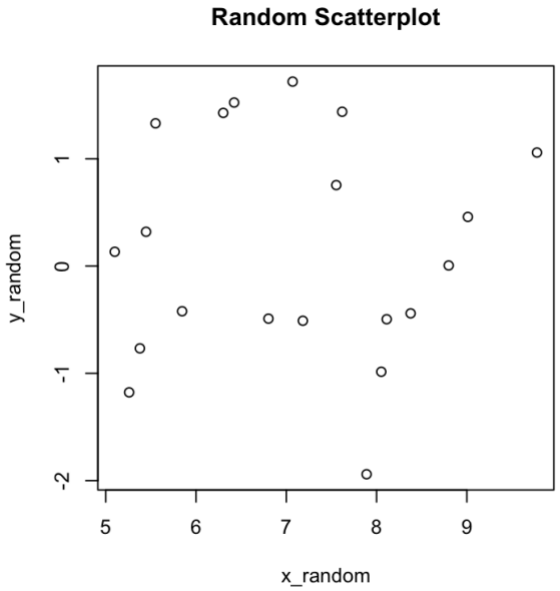
Q8 Density Figure File

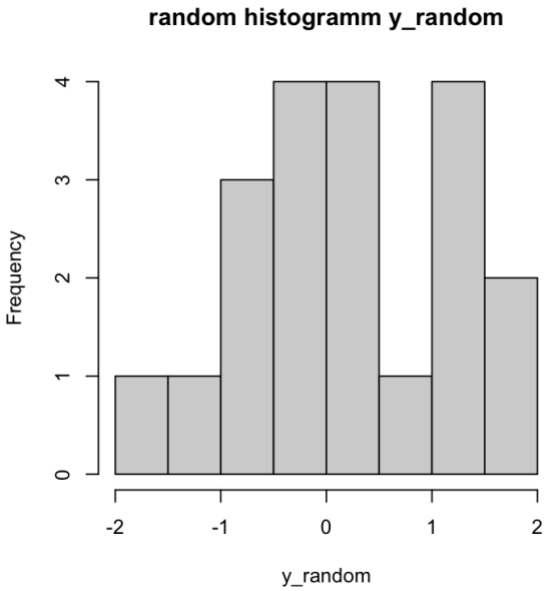


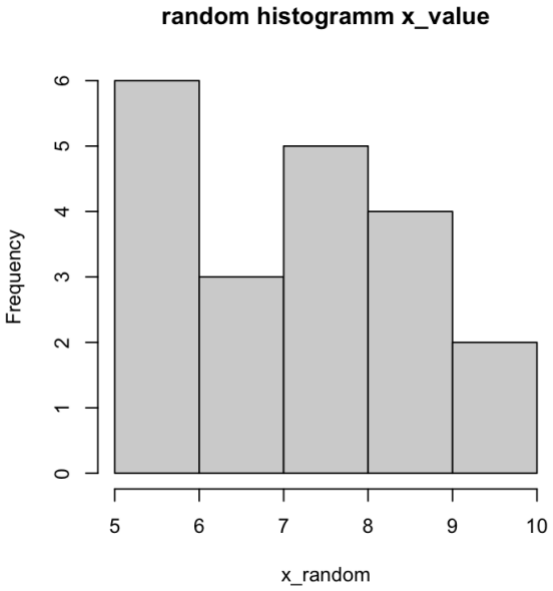
Q9 Random Data Set

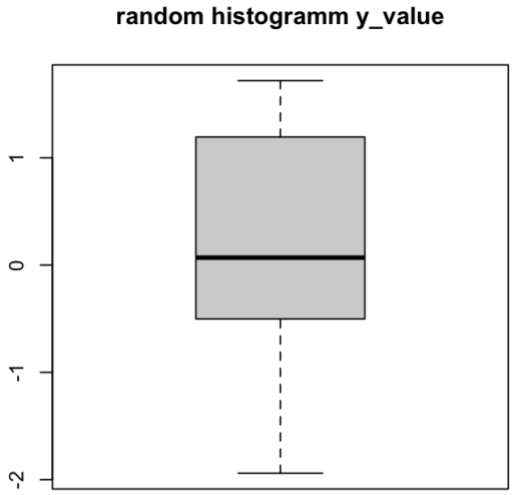
```
set.seed(36)
n_pts = 20
x_min = 5
x_max = 10

# X values are uniformly distributed
x_random = runif(n = n_pts, min = x_min, max = x_max)

# Y values are normally-distributed.
# default parameters for mean and sd.
y_random = rnorm(n = n_pts)

dat_random = data.frame(x = x_random, y = y_random)

par(mfrow= c(2,2))

plot(x = x_random, y = y_random, main = "Random Scatterplot")
```

Q10 Random Data Image File





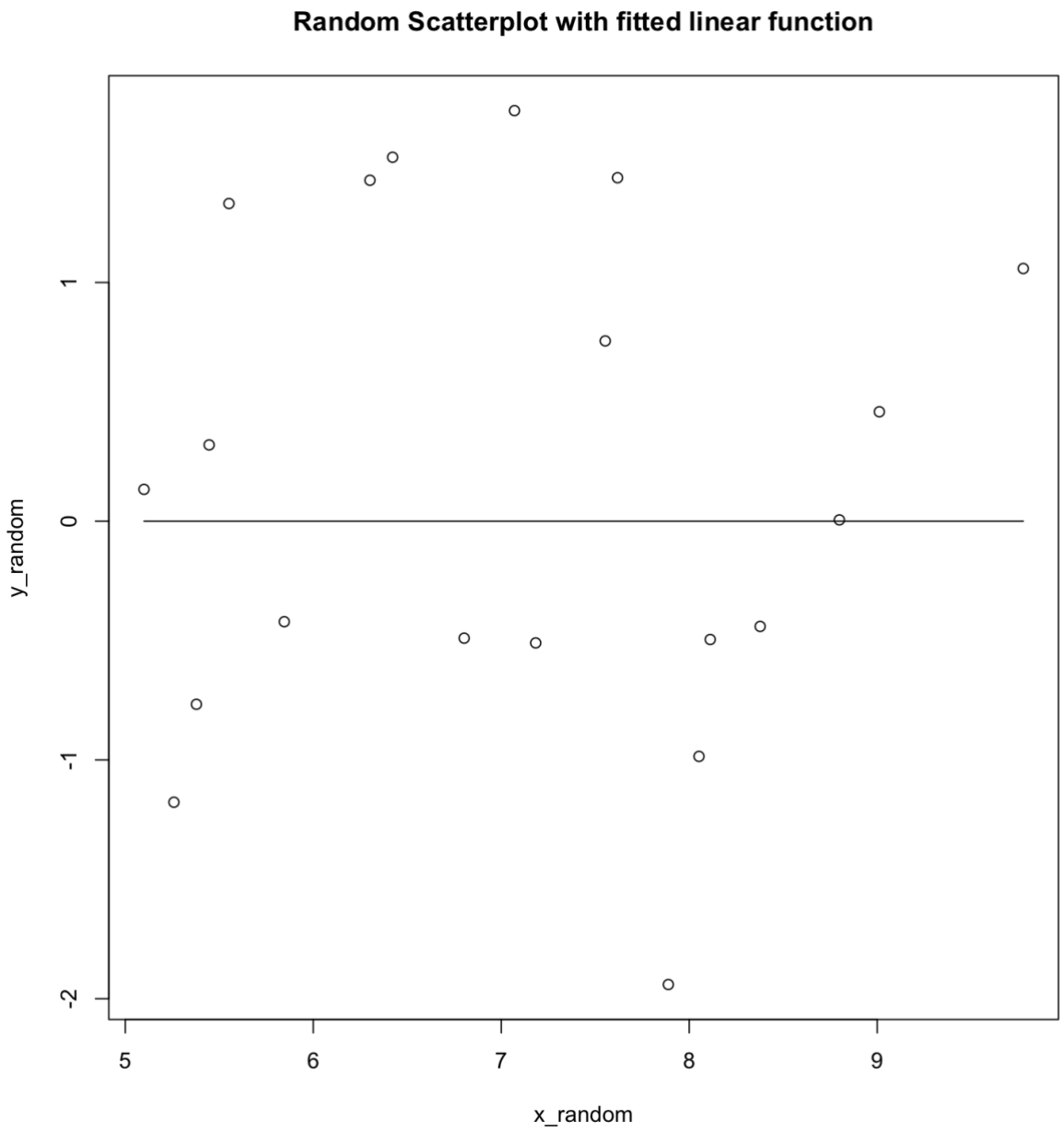




Q11 Random Data Model Fit

```
require(here)
png(
  filename = here("plot with line"),
  width = 1500, height = 1600,
  res = 180, units = "px")

plot(x = x_random, y = y_random, main = "Random Scatterplot with fitted linear function")

guess_x = 7.5
guess_y = 0
guess_slope = 0


curve(line_point_slope(x, guess_x, guess_y, guess_slope), add = T)
dev.off() #save file
```

Q12 Random Data Model Fit Image File



Q13 Random Data Model Residuals

```
dat_random$y_predicted <- line_point_slope(dat_random$x, guess_x, guess_y, guess_slope)
head(dat_random)
```

```
dat_random$resids = y_random - dat_random$y_predicted
head(dat_random)
sum(dat_random$resids)
```

Q14 Random Data Model Residual Plot