Julian Burgoff

10/10/2022

Analysis of Environmental Data Lab

Lab 4

1. n\_samples\_1= 17

n\_samples\_2= 30

n\_samples\_3= 300

n\_samples\_4= 3000

pop\_sd = 2.4

pop\_mean = 10.4

dat\_1 = rnorm(n = n\_samples\_1, mean = pop\_mean, sd = pop\_sd)

dat\_2 = rnorm(n = n\_samples\_2, mean = pop\_mean, sd = pop\_sd)

dat\_3 = rnorm(n = n\_samples\_3, mean = pop\_mean, sd = pop\_sd)

dat\_4 = rnorm(n = n\_samples\_4, mean = pop\_mean, sd = pop\_sd)

1. require(here)

png(

filename = here("lab\_04\_hist\_01.png"),

width = 1500, height = 1600,

res = 180, units = "px")

par(mfrow = c(2, 2))

hist(dat\_1, main= "Histogram of 17 randomly generated numbers")

hist(dat\_2, main= "Histogram of 30 randomly generated numbers")

hist(dat\_3, main= "HIstogram of 300 randomly generated numbers")

hist(dat\_4, main= "Histogram of 3000 randomly generated numbers")

dev.off()

1. PNG file attached
2. The histogram of 17 randomly generated numbers is relatively normally distributed, minus the big gap where the frequency should be highest around 11. The histogram of 30 numbers is skewed to the right and more uniformly distributed. The histogram of 300 numbers is much more normally distributed and the histogram of 3000 is almost perfectly normally distributed, where the histogram peaks at the mean and falls evenly on each side.
3. The shapes are different because of the different sample sizes represented in each histogram. As the sample size goes up, the random numbers get closer and closer to a normal distribution about the mean.
4. The parameters are the mean and standard deviation (mean=0, sd=1).
5. require(here)

svg(

filename = here("norm\_1.svg"),

width = 10, height = 10)

x = seq(-100, 100, length.out = 1000)

y = dnorm(x, mean= 10.4, sd= 2.4)

plot(x, y, main = "Normal PDF with Mean = 10.4, SD = 2.4", type = "l", xlim = c(3, 19))

abline(h = 0)

dev.off()

1. norm\_1.svg attached
2. n\_pts = 111

x\_min = 0

x\_max = 10

set.seed(1)

dat\_unif\_1 = runif(n = n\_pts, min = x\_min, max = x\_max)

1. quad\_plot.svg attached
2. n\_pts = 111

x\_min = 0

x\_max = 10

x\_random = runif(n = n\_pts, min = x\_min, max = x\_max)

y\_random = rnorm(n = n\_pts)

dat\_random = data.frame(x = x\_random, y = y\_random)

1. lin\_function.svg attached
2. y\_predicted= line\_point\_slope(dat\_random$x, guess\_x, guess\_y, guess\_slope)

dat\_random$y\_predicted<- y\_predicted

resids= dat\_random$y\_predicted- dat\_random$y

dat\_random$resids<- resids

1. Chart, histogram

   Description automatically generated

Chart, scatter chart

Description automatically generated