> ### CS 544 Module 4 Assignment

> ### Hao Wu

> ### Part 1) Binomial distribution

> n <- 5

> p <- 0.4

> ## Question a

> # PMF Plot and Caculate

> pmf <- dbinom(0:n, size = n, prob = p)

> plot(0:n, pmf, type = "h", xaxt = "n",

+ main = "", xlab = "x", ylab = "PMF")

> points(0:n, pmf, pch = 16)

> axis(side = 1, at = 0:n, labels = TRUE)

> abline(h = 0, col="red")

> #CDF Plot and caculate

> cdf <- pbinom(0:n, size = n, prob = p)

> cdf <- c(0, cdf)

> cdfplot <- stepfun(0:n, cdf)

> plot(cdfplot, verticals = FALSE, pch = 16,

+ main = "", xlab = "x", ylab = "CDF")

> ## Question b

> Qb <- dbinom(2, size = n, prob = p)

> Qb

[1] 0.3456

> # or

> choose(n,2) \* p^2 \* (1 - p)^3

[1] 0.3456

> ## Question c

> Qc <- sum(dbinom(2:n, size = n, prob = p))

> Qc

[1] 0.66304

> ## Question d

> y <- rbinom(1000, size=5, prob=p)

> table(y)

y

0 1 2 3 4 5

80 246 345 257 63 9

> plot(table(y), type="h", col="red")

> ### Part 2) Negative Binomial distribution

> r <- 3

> p <- 0.6

> ## Question a

> pmf <- dnbinom(0:10, size = r, prob = p)

> plot(0:10,pmf,type="h",

+ xlab="x",ylab="PMF", ylim = c(0, 0.2))

> abline(h=0, col="red")

> cdf <- c(0, cumsum(pmf))

> cdfplot <- stepfun(0:10, cdf)

> plot(cdfplot, verticals = FALSE, pch = 16,

+ main = "", xlab = "x", ylab = "CDF")

> ## Question b

> dnbinom(4, size = r, prob = p)

[1] 0.082944

> # or

> choose(6, 2) \* (p\*\*r) \* ((1-p)\*\*4)

[1] 0.082944

> ## Question c

> pnbinom(4, size = r, prob = p)

[1] 0.903744

> # or

> sum(dnbinom(0:4, size = r, prob = p))

[1] 0.903744

> ## Question d

> x <- rnbinom(1000, size = r, prob = p)

> plot(table(x))

> M <- 60 ## Mult Choice Question

> N <- 40 ## Programming Question

> K <- 20 ## The question which will be select on final exam

> ## Question a

> # PMF

> pmf <- dhyper(0:K, m = M, n = N, k = K)

> plot(0:K,pmf,type="h",

+ xlab="x",ylab="PMF", ylim = c(0, 0.2))

> abline(h=0, col="red")

> # CDF

> cdf <- c(0, cumsum(pmf))

> cdfplot <- stepfun(0:K, cdf)

> plot(cdfplot, verticals = FALSE, pch = 16,

+ main = "", xlab = "x", ylab = "CDF")

> ## Question b

> dhyper(10, m = M, n = N, k = K)

[1] 0.1192361

> # or

> choose(M,10) \* choose(N,10) / choose(M+N, 20)

[1] 0.1192361

> ## Question c

> 1 - phyper(10, m = M, n = N, k = K)

[1] 0.77902

> # or

> phyper(10, m = M, n = N, k = K, lower.tail = FALSE)

[1] 0.77902

> ## Question d

> z <- rhyper(1000, m = M, n = N, k = K )

> plot(table(z))

> ## Question a

> dpois(8, lambda=10)

[1] 0.112599

> ## Question b

> sum(dpois(0:8, lambda=10))

[1] 0.3328197

> #or

> ppois(8, lambda=10)

[1] 0.3328197

> ## Question c

> sum(dpois(6:12, lambda=10))

[1] 0.7244705

> #or

> ppois(12, lambda=10) - ppois(5, lambda=10)

[1] 0.7244705

> #or

> diff(ppois(c(5,12), lambda=10))

[1] 0.7244705

> # Plot PMF

> pmf <- dpois(0:20, lambda=10)

> plot(0:20,pmf,type="h",

+ xlab="x",ylab="PMF", ylim = c(0, 0.25))

> abline(h=0, col="red")

> # Plot CDF

> cdf <- c(0, cumsum(pmf))

> cdfplot <- stepfun(0:20, cdf)

> plot(cdfplot, verticals = FALSE, pch = 16,

+ main = "", xlab = "x", ylab = "CDF")

> ## Question e

> x <- rpois(50, lambda = 10)

> plot(table(x))

> ### Part 5) Normal Distribution

> mu <- 100

> sd <- 10

> ## Question a

> # Plot PDF

> x <- seq(mu-3\*sd, mu+3\*sd)

> pdf = dnorm(x,mean = mu,sd = sd)

> plot(x, pdf, type="l", col="red",

+ xlim=c(mu-3\*sd,mu+3\*sd), ylim=c(0,0.05),

+ xaxt="n", yaxt="n",

+ main="Normal Distribution", xlab="probability", ylab="sd")

> axis(side = 1, at = c(70,80,90,100,110,120,130),

+ labels = TRUE)

> axis(side = 2, at = c(0,0.01,0.02,0.03,0.04,0.05),

+ labels = TRUE)

> ## Question b

> pnorm(120, mean = mu, sd = sd,lower.tail = FALSE)

[1] 0.02275013

> #or

> 1- pnorm(120,mean = mu, sd = sd )

[1] 0.02275013

> ## Question c

> pnorm(90, mean = mu, sd= sd) - pnorm(80, mean= mu, sd=sd)

[1] 0.1359051

> ## Question d

> #within three standard deviation

> pnorm(mu + 3\*sd, mean = mu, sd = sd) -

+ pnorm(mu - 3\*sd, mean = mu, sd = sd)

[1] 0.9973002

> #within two standard deviation

> pnorm(mu + 2\*sd, mean = mu, sd = sd) -

+ pnorm(mu - 2\*sd, mean = mu, sd = sd)

[1] 0.9544997

> #within one standard deviation

> pnorm(mu + sd, mean = mu, sd = sd) -

+ pnorm(mu - sd, mean = mu, sd = sd)

[1] 0.6826895

> ## Question e

> qnorm(0.1, mean= mu , sd= sd)

[1] 87.18448

> qnorm(0.9, mean= mu , sd= sd)

[1] 112.8155

> ## Question f

> qnorm(0.98, mean = mu, sd= sd)

[1] 120.5375

> ## Question g

> x <- rnorm(10000, mean = mu, sd = sd)

> x <- round(x)

> plot(table(x))