**Problem 2 (Probability Distributions in R)**

Set your seed (pick a number)

**set.seed(5678)**

1. Create a vector `x' containing 10000 random samples from a chi-square distribution with 10 degrees of freedom (hint:?rchisq)

**x <- rchisq(n=10000, df=10)**

**head(x)**

**#Answer - [1] 7.795379 7.668076 5.437041 15.274104 19.933838 5.029184**

1. Find the value of the chi-square distribution with 10 degrees of freedom corresponding to the 99th percentile (hint: ?qchisq)

**y <- qchisq(p = .99, df = 10)**

**y**

**#Answer - [1] 23.20925**

1. What percentage of observations in your sample fall above this value? Is it exactly 1%? Should it be? Discuss.

**ct <- length(which(x > y))**

**ct #Answer - [1] 102**

**(ct / length(x > y)) \* 100 #Answer - [1]**

**Since I am using random chisq, there is a chance that the percentile won't be exactly 1%**

**because of the degree of freedom.**

1. Using ggplot, display a density histogram of the vector 'x' (with 50 bins and your choice of colors for the boundaries and fill) and overlay the histogram with the density plot of the vector (size of the curve should greater than or equal to 1.5 and adjust value should be 1.5).

**library(ggplot2)**

**library(dplyr)**

**a1 <- ggplot() +**

**geom\_histogram(aes(x), bins = 50, size = 1.5)**

**fill <- c("blue")**

**a1 <- a1 + labs(x="x Values", y="y values") + ggtitle("Histogram with density plot")**

**a1**

**ggplot() + geom\_histogram(aes(x), bins=50,size=1.5)**