2d Goodness of Fit

Karan Vijay Singh 20745105 STAT 847

H_0 is that all of the following hold simultaneously:

- X and Y are independently distributed random variates, with
- $X \sim U(0,13)$ and
- $Y \sim U(0,8)$.

1. **(15 marks)** Carry out a goodness of fit test of H_0 and report the evidence against it. Do this by dividing the two page spread into three by three grid of 9 equal area rectangles. Show **all** of your code and work.

```
#setting the path for the data
dataDirectory <- "C:/Users/Jasdeep/Desktop/SEM 2 Waterloo/STAT
847/Assignments/dataDirectory"
path_concat <- function(path1, path2, sep="/") paste(path1, path2, sep = sep)

#loading the waldo data
load(path_concat(dataDirectory, "waldo.Rda"))

x <- waldo$x
y <- waldo$y

observedValues <- c()
binbreaks_y <- seq(0, ceiling(max(y)), length.out =4)
binbreaks_x <- seq(0, ceiling(max(x)), length.out =4)
Bx <- (length(binbreaks_x)-1)
By <- (length(binbreaks_y)-1)


#getting observed values for each of the bins
getObservedValues <- function(sampled_data){
  x <- sampled_data$x
  y <- sampled_data$y
  for(i in 1:Bx){
    for(j in 1:By){
      observedValues <- c(observedValues,sum(binbreaks_y[j] < y & y <=
binbreaks_y[j+1] & binbreaks_x[i] < x & x <= binbreaks_x[i+1]))
    }
  }
  return (observedValues)
}


actualValues <- getObservedValues(waldo)
#getting expected values for each of the bins
```

```r
#in case of uniform it will be same number in each bin
B <- length(actualValues)
probs <- rep(1/B, B)
actual_pValue <- chisq.test(actualValues, p = probs)
actual_pValue

##
##  Chi-squared test for given probabilities
##
## data:  actualValues
## X-squared = 13, df = 8, p-value = 0.1118
```

The p value is 0.1118 (small), hence we can say that we have got some evidence against the null hypothesis .