STATS 302 - Assignment 3

Jonathan Kumarich

18/05/2021

### Packages and clear

rm(list=ls())  
  
pacman::p\_load(  
 ggplot2,  
 rstudioapi,  
 tinytex,  
 dplyr  
)

### Data Import

setwd(dirname(rstudioapi::getActiveDocumentContext()$path))  
getwd()

## [1] "/Users/jonathankumarich/Uni/Uni/STATS 302/Assignments/A3/STATS302-A3"

fungal <- read.csv("fungal-all.csv")  
fungal.desc <- fungal[, 1:5]  
fungal.val <- fungal[ ,-c(1:5)]  
rm(fungal)

### Question 1

pvals <- rep(NA, ncol(fungal.val))  
  
for (i in 1:ncol(fungal.val)) {  
 test <- kruskal.test(fungal.val[, i] ~ fungal.desc$Condition)  
 pvals[i] <- test$p.value  
}  
  
ggplot(data.frame(pvals), aes(x=pvals)) +   
 geom\_histogram(fill="#004775", color="black") +   
 theme\_classic() +  
 xlab("P-Values") +   
 ylab("Count") +   
 ggtitle("Distribution of P-Values") +  
 theme(plot.title = element\_text(hjust = 0.5))

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

