

```
In [1]: ## Data prep

M2 <- matrix(c(6,3,10,13) ,2,2)
M3 <- matrix(c(7,3,4,7) ,2,2)

M <- rbind(M2, M3)
M <- cbind(rep(c(2, 3), each = 2), M)

colnames(M) <- c("Z", "Yes", "No")

df <- as.data.frame(M)
```

```
In [3]: ## CMH statistic

cmh <- function(x) {
  require(tidyverse)
  Z_2 <- c(2, 3)
  mu_n11k <- function(y) apply(y, 1, sum)[1] *
    apply(y, 2, sum)[1] / sum(y)
  deviations_n11k <- function(y) (y[1,1]- mu_n11k(y))
  variances_n11k <- function(y) prod(apply(y, 1, sum)) *
    prod(apply(y, 2, sum)) / (sum(y)^2 *(sum(y) - 1))

  temp1 <-0
  temp2 <-0

  for(k in 1:length(Z_2)){
    temp1 <- temp1 + x %>% filter(Z == Z_2[k]) %>%
      select(!Z) %>% deviations_n11k
    temp2 <- temp2 + x %>% filter(Z == Z_2[k]) %>%
      select(!Z) %>% variances_n11k
  }

  result <- temp1^2/temp2
  names(result) <- "CMH statistic"

  return(result)
}

cmh(df)
cmh(df) > qchisq(0.05, df=1, lower.tail = F)
```

CMH statistic: 3.49858394415148

CMH statistic: FALSE