1)A.

Estimation of Odds ratio for Birds Vs No Birds Groups = 2.257145

Odds of getting a cancer in Birds group are estimated to be 2.257145 times higher than the odds of getting cancer in No Birds group.

B)

odds ratio with 95% C.I. estimate lower upper

Bird 1.000000 NA NA

No Bird 2.257145 1.60518 3.173915

95% confidence interval for no birds group is (1.60518, 3.173915). As the CI does not include 1, we can reject the null hypothesis that the probability of cancer for both the Birds and No birds group are the same.  
  
For this, we can conclude with 95% confidence that there is a relationship between bird ownership and lung cancer.

2) A)

Odds ratio of Trt Vs Ctrl for Study 1, 0.19519

Odds ratio of Trt Vs Ctrl for Study 2, 1.01210

Odds ratio of Trt Vs Ctrl for Study 3, 0.62391

B)

Breslow-Day p-value = 0.0001456754 < 0.001 , we can conclude that the odds ratio are different across the 3 studies .

C)

After running CMH Test using mantelhaen.test, we get

Common Odds Ratio = 0.9570013, p-value = 0.4663.

Conclusion: As 95% CI includes 1, we can not reject the hypothesis of the average of odds ratio across different studies equals to 1.

3)A)

Sample mean , 2.104167

B)

GOF test statistic = 6.340026

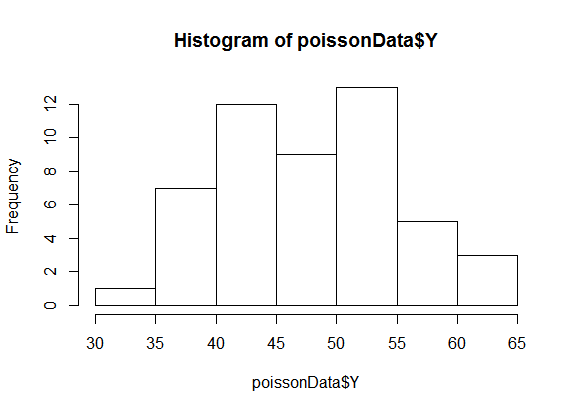
p-value = 0.09618956 > 0.05

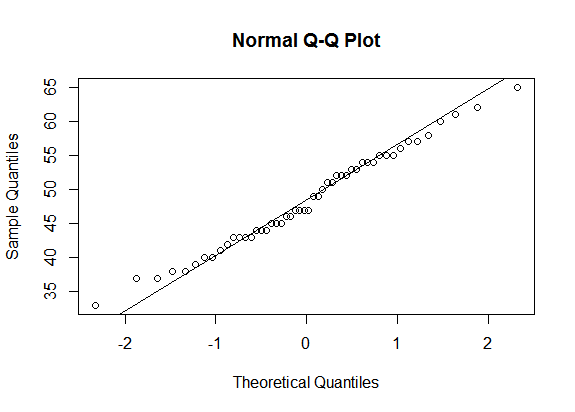
Conclusion: No evidence against the poisson assumption.

4)A)

Sample mean = 48.38

Sample standard deviation = 7.342607





B)

Confidence interval = (46.29325 50.46675)

C)

Confidence interval = (46.45201,50.30799)

|  |
| --- |
| > #Q1  > birds <- matrix(c(98,101,141,328),nrow = 2,byrow = TRUE)  > rownames(birds) <- c("Bird","No Bird")  > colnames(birds) <- c("Cancer Patients","Healthy Controls")  > birds  Cancer Patients Healthy Controls  Bird 98 101  No Bird 141 328  > library(epitools)  > oddsratio(birds,method = "wald")  $data  Cancer Patients Healthy Controls Total  Bird 98 101 199  No Bird 141 328 469  Total 239 429 668  $measure  NA  odds ratio with 95% C.I. estimate lower upper  Bird 1.000000 NA NA  No Bird 2.257145 1.60518 3.173915  $p.value  NA  two-sided midp.exact fisher.exact chi.square  Bird NA NA NA  No Bird 3.052348e-06 3.938413e-06 2.243712e-06  $correction  [1] FALSE  attr(,"method")  [1] "Unconditional MLE & normal approximation (Wald) CI"  >  > #Q2  > BCG\_data <- array(c(8,10,2537,619,  + 505,499,87886,87892,  + 29,45,7470,7232),  + dim = c(2,2,3),  + dimnames = list(Trt=c("Trt","Ctrl"),  + Response = c("TBpos","TBneg"),  + Study = c("1","2","3")))  > BCG\_data  , , Study = 1  Response  Trt TBpos TBneg  Trt 8 2537  Ctrl 10 619  , , Study = 2  Response  Trt TBpos TBneg  Trt 505 87886  Ctrl 499 87892  , , Study = 3  Response  Trt TBpos TBneg  Trt 29 7470  Ctrl 45 7232  >  > #Response by study  > Trt.by.study <- margin.table(BCG\_data,c(1,3))  > barplot(Trt.by.study,legend=T,main = "Treatment By Study")  > resp.by.study <- margin.table(BCG\_data,c(2,3))  > barplot(resp.by.study,legend=T,main = "Response By Study")  >  > library(lawstat)  > cmh.test(BCG\_data)  Cochran-Mantel-Haenszel Chi-square Test  data: BCG\_data  CMH statistic = 0.53072, df = 1.00000, p-value = 0.46631, MH Estimate = 0.95700, Pooled  Odd Ratio = 0.95685, Odd Ratio of level 1 = 0.19519, Odd Ratio of level 2 = 1.01210, Odd  Ratio of level 3 = 0.62391  >  >  > #B  > library(metafor)  > cmh <- rma.mh(ai = BCG\_data[1,1,],bi= BCG\_data[1,2,],ci= BCG\_data[2,1,], di = BCG\_data[2,2,])  > cmh  Fixed-Effects Model (k = 3)  Test for Heterogeneity:  Q(df = 2) = 15.1279, p-val = 0.0005  Model Results (log scale):  estimate se zval pval ci.lb ci.ub  -0.0440 0.0605 -0.7259 0.4679 -0.1626 0.0747  Model Results (OR scale):  estimate ci.lb ci.ub  0.9570 0.8499 1.0776  Cochran-Mantel-Haenszel Test: CMH = 0.4874, df = 1, p-val = 0.4851  Tarone's Test for Heterogeneity: X^2 = 17.6682, df = 2, p-val = 0.0001  >  > #Breslow-Day test  > cmh$BD  [1] 17.66826  > cmh$BDp  [1] 0.0001456754  >  > #c  > mantelhaen.test(BCG\_data,correct = FALSE)  Mantel-Haenszel chi-squared test without continuity correction  data: BCG\_data  Mantel-Haenszel X-squared = 0.53072, df = 1, p-value = 0.4663  alternative hypothesis: true common odds ratio is not equal to 1  95 percent confidence interval:  0.8499098 1.0775866  sample estimates:  common odds ratio  0.9570013  >  >  >  > #3  > #Observed data  >  > Obs <- c(109,65,22,3,1)  > Y <- seq(from=0,to=4,by=1)  > Y  [1] 0 1 2 3 4  >  > #Calculate the mean  > Muhat <- sum(Obs\*Y)/sum(Obs)  > Muhat  [1] 0.61  >  > #Calculate the corresponding Poisson Probabilities  > Prob <- dpois(Y,Muhat)  > Prob  [1] 0.543350869 0.331444030 0.101090429 0.020555054 0.003134646  >  > length(Prob)  [1] 5  > sum(Prob)  [1] 0.999575  >  > #Fix the final entry so that the probabilities sum to 1  > Prob[5] <- 1-ppois(3,Muhat)  > Prob  [1] 0.543350869 0.331444030 0.101090429 0.020555054 0.003559618  > sum(Prob)  [1] 1  > Prob  [1] 0.543350869 0.331444030 0.101090429 0.020555054 0.003559618  >  > #Calculate Expected values and Contributions to Chisquare TS  > Exp <- Prob\*200  > X2 <- (Obs-Exp)^2/Exp  > cbind(Y,Obs,Prob,Exp,X2)  Y Obs Prob Exp X2  [1,] 0 109 0.543350869 108.6701738 0.00100106  [2,] 1 65 0.331444030 66.2888060 0.02505734  [3,] 2 22 0.101090429 20.2180858 0.15704840  [4,] 3 3 0.020555054 4.1110108 0.30025340  [5,] 4 1 0.003559618 0.7119235 0.11656877  >  > #Run GOF test  > ChisqTS <- sum(X2)  > ChisqTS  [1] 0.599929  > pval <- 1-pchisq(ChisqTS,5-2)  > pval  [1] 0.8964486  >  > #4  > poissonData <- read.csv("C:\\Users\\SHAIKHSHAWON\\Dropbox\\Fall 2015\\Stat 511\\Homework\\Homework10\\PoissonData.csv")  > poissonData  Y  1 43  2 45  3 47  4 45  5 55  6 58  7 41  8 43  9 40  10 51  11 46  12 47  13 55  14 49  15 44  16 53  17 33  18 65  19 43  20 50  21 38  22 57  23 37  24 51  25 54  26 54  27 56  28 44  29 61  30 43  31 46  32 57  33 47  34 52  35 38  36 47  37 39  38 37  39 44  40 49  41 53  42 45  43 62  44 60  45 40  46 52  47 54  48 42  49 55  50 52  > str(poissonData)  'data.frame': 50 obs. of 1 variable:  $ Y: int 43 45 47 45 55 58 41 43 40 51 ...  > mean(poissonData$Y)  [1] 48.38  > s<-sd(poissonData$Y)  > s  [1] 7.342607  > s^2  [1] 53.91388  > hist(poissonData$Y)  > qqnorm(poissonData$Y)  > qqline(poissonData$Y)  > t.test(poissonData$Y)  One Sample t-test  data: poissonData$Y  t = 46.591, df = 49, p-value < 2.2e-16  alternative hypothesis: true mean is not equal to 0  95 percent confidence interval:  46.29325 50.46675  sample estimates:  mean of x  48.38  >  > y <- sum(poissonData$Y)  > y  [1] 2419  >  > LowerCI <- y - (1.96\*sqrt(y))  > UpperCI <- y + (1.96\*sqrt(y))  > LowerCI/50  [1] 46.45201  > UpperCI/50  [1] 50.30799 |
|  |
| |  | | --- | | > | |