Data Science HW6 Total: 20 points Due: February 15th, 2019, NOON Assignment Paper: Shaw et al., 2016, Shaw S1 S2

1. In your own words, specify the research question(s) (2 sentences or less) for Study 2 and Study S2

* Study 2: Are older children more willing to accept self-produced disadvantageous inequality than disadvantageous inequality made by others?
* S2: Do children become more generous as they get older as measured by whether they allocate an extra item to another person, regardless of whether they will be disadvantaged? Said another way, are childrens’ views of inequity influenced by whether the resource is to be shared between themselves and another or between two others.

1. Include and explain IVs, DVs for Study 2 and Study S2

* Study 2:
  + IV: Agegroup (older vs younger), condition/agent (disadvantaged by others, disadvantaged by self)
  + DV: choice
* S2:
  + IV: Age (agegroup: older vs younger), condition (disadvantaging vs 3rd party)
  + DV: choice

1. R section (please complete the following and include your script and outputs as in-line text below) Note: For Studies 2 & S2, the authors chose to rescale Agegroup so young: 0.50, old: -0.5). By rescaling, Agegroup is now treated as a continuous variable. Be sure to perform the necessary transformations before running the regression
   1. Reproduce Study 2’s logistic regression analyses (include regression output (beta, SE, etc.) and all Wald statistics)
   2. Reproduce Study S2’s logistic regression analyses (include regression output (beta, SE, etc.) and all Wald statistics) [6]

hw\_06

Abby Bergman

2/13/2019

#rescale agegroup  
study2 = study2 %>%  
 select(-X, -X.1, -X.2, -X.3, -X.4, -X.5, -X.6, -X.7)  
  
colnames(study2) = c("DV", "condition", "agegroup", "age", "sex")  
  
 study2 <- study2%>%  
 mutate(agegroup = ifelse(study2$agegroup == 1, .5, -.5)) %>%  
 mutate(DV = as.factor(DV))  
  
colnames(s2) = c("DV", "condition", "agegroup", "age", "sex")  
  
s2 <- s2%>%  
 mutate(agegroup = ifelse(s2$agegroup == 1, .5, -.5))

#a. Reproduce Study 2’s logistic regression analyses (include regression output (beta, SE, etc.) and all Wald statistics)   
  
fitstudy2 = glm(DV ~ agegroup \* condition, data = study2, family = binomial)  
  
summary(fitstudy2)

##   
## Call:  
## glm(formula = DV ~ agegroup \* condition, family = binomial, data = study2)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.3153 -0.9282 -0.7352 1.0455 1.6973   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -0.6129 0.1617 -3.790 0.000151 \*\*\*  
## agegroup -0.3741 0.3234 -1.157 0.247404   
## condition -0.9252 0.3234 -2.860 0.004232 \*\*   
## agegroup:condition 1.1267 0.6469 1.742 0.081551 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 247.51 on 190 degrees of freedom  
## Residual deviance: 235.48 on 187 degrees of freedom  
## (29 observations deleted due to missingness)  
## AIC: 243.48  
##   
## Number of Fisher Scoring iterations: 4

Anova(fitstudy2, type = 3, test.statistic = "Wald")

## Analysis of Deviance Table (Type III tests)  
##   
## Response: DV  
## Df Chisq Pr(>Chisq)   
## (Intercept) 1 14.3612 0.0001509 \*\*\*  
## agegroup 1 1.3379 0.2474038   
## condition 1 8.1813 0.0042325 \*\*   
## agegroup:condition 1 3.0337 0.0815513 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Reproduce Study S2’s logistic regression analyses (include regression output (beta, SE, etc.) and all Wald statistics) [6]  
  
  
fits2 = glm(DV ~ agegroup \* condition, data = s2, family = binomial)  
  
summary(fits2)

##   
## Call:  
## glm(formula = DV ~ agegroup \* condition, family = binomial, data = s2)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.2388 -0.8369 -0.6476 1.1173 2.5674   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.3598 0.3028 -4.492 7.07e-06 \*\*\*  
## agegroup 0.3953 0.6055 0.653 0.513818   
## condition -1.4075 0.6055 -2.324 0.020100 \*   
## agegroup:condition 3.9874 1.2110 3.293 0.000992 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 157.19 on 135 degrees of freedom  
## Residual deviance: 136.53 on 132 degrees of freedom  
## AIC: 144.53  
##   
## Number of Fisher Scoring iterations: 5

Anova(fits2, type = 3, test.statistic = "Wald")

## Analysis of Deviance Table (Type III tests)  
##   
## Response: DV  
## Df Chisq Pr(>Chisq)   
## (Intercept) 1 20.1743 7.07e-06 \*\*\*  
## agegroup 1 0.4263 0.5138178   
## condition 1 5.4032 0.0201001 \*   
## agegroup:condition 1 10.8417 0.0009924 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1