Vignettes for `additive_interactions`

Maya B. Mathur & Tyler J. VanderWeele

March 8, 2017

Example 1: Exposures positively associated with outcome

Simulate data with two exposures (`E1` and `E2`) and confounder `C`:

```
n=5000
set.seed(451)
E1 = rbinom(n, size=1, p=0.2)   # generate a binary genetic exposure
set.seed(1984)
E2 = rbinom(n, size=1, p=.75)

# binary confounder that is correlated with both exposures
# linear predictor from which to generate C
linpred = 1 + 1.5*E1 + 2.2*E2
# expit to back-transform logits into probabilities
probs = exp(linpred) / (1 + exp(linpred))
# generate C
set.seed(451)
C = rbinom(n, size=1, p=probs)

# true betas for Y
beta0 = log(1)
betaE1 = log(1.17)
betaE2 = log(1.15)
betaE1E2 = log(1.7)

# for generation of Y
linpred = beta0 + betaE1*E1 + betaE2*E2 + betaE1E2*E1*E2
probs = exp(linpred) / (1 + exp(linpred))
set.seed(22)
Y = rbinom(n, size=1, p=probs)
```

First fit the logistic regression model including firstly the main effects and interaction of the exposures of interest, followed by any covariates:

```
m = glm(Y ~ E1*E2 + C, family=binomial(link="logit"))
summary(m)
```

```
##
## Call:
## glm(formula = Y ~ E1 * E2 + C, family = binomial(link = "logit"))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6143  -1.2324   0.7963   1.1234   1.1635
##
## Coefficients:
```

1

```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.03282    0.16088   0.204 0.838365
## E1           0.23828    0.14837   1.606 0.108279
## E2           0.07668    0.07428   1.032 0.301918
## C            0.01890    0.16113   0.117 0.906634
## E1:E2        0.61931    0.17272   3.586 0.000336 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 6863.1  on 4999  degrees of freedom
## Residual deviance: 6757.5  on 4995  degrees of freedom
## AIC: 6767.5
##
## Number of Fisher Scoring iterations: 4
```

Then pass it to `additive_interactions`:

```
additive_interactions(m)
```

```
## Loading required package: msm
```

```
##     Stat        Est       CI.lo      CI.hi      p.val.0 p.val.epi
## 1  RERI 1.19661040  0.65430493 1.7389159 7.636672e-06 0.9981552
## 2    AP 0.47011133  0.30355991 0.6366628 1.580897e-08        NA
## 3    E1 0.17411092 -0.05795168 0.4061735 7.071139e-02        NA
## 4    E2 0.05157272 -0.04212606 0.1452715 1.403422e-01        NA
## 5 E1:E2 0.77431636  0.50702237 1.0416103 6.823338e-09        NA
##   p.val.suff.cause
## 1        0.2386735
## 2               NA
## 3               NA
## 4               NA
## 5               NA
```

We can see how inference on the mechanistic interactions becomes less stringent if we assume monotonicity for at least one of the exposures:

```
additive_interactions(m, monotone=1)
```

```
##
## Assuming AT LEAST ONE of exposures has positive monotonic effect
##
##     Stat        Est       CI.lo      CI.hi      p.val.0 p.val.epi
## 1  RERI 1.19661040  0.65430493 1.7389159 7.636672e-06 0.2386735
## 2    AP 0.47011133  0.30355991 0.6366628 1.580897e-08        NA
## 3    E1 0.17411092 -0.05795168 0.4061735 7.071139e-02        NA
## 4    E2 0.05157272 -0.04212606 0.1452715 1.403422e-01        NA
## 5 E1:E2 0.77431636  0.50702237 1.0416103 6.823338e-09        NA
##   p.val.suff.cause
## 1        0.2386735
```

```
## 2              NA
## 3              NA
## 4              NA
## 5              NA
```

Example 2: Exposures negatively associated with outcome

Here, the exposure and interaction effects are negatively associated with the outcome, so the function will automatically recode the exposures against new reference levels:

```
# preventive case
# true betas
beta0 = log(1)
betaE1 = log(0.95)
betaE2 = log(0.9)
betaE1E2 = log(0.67)

# for generation of Y
linpred = beta0 + betaE1*E1 + betaE2*E2 + betaE1E2*E1*E2
probs = exp(linpred) / (1 + exp(linpred))
set.seed(22)
Y = rbinom(n, size=1, p=probs)

# put in dataframe since the function will need to recode exposures
d = data.frame(E1, E2, Y)
```

Again, begin by fitting the logistic regression model:

```
m = glm(Y ~ E1*E2 + C, family=binomial(link="logit"), data=d)
summary(m)
```

```
##
## Call:
## glm(formula = Y ~ E1 * E2 + C, family = binomial(link = "logit"),
##     data = d)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.2951  -1.1389  -0.9147   1.2164   1.4652
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.27243    0.15955   1.708   0.0877 .
## E1          -0.30693    0.14806  -2.073   0.0382 *
## E2          -0.12166    0.07423  -1.639   0.1012
## C           -0.24205    0.15930  -1.519   0.1286
## E1:E2       -0.25683    0.16988  -1.512   0.1306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 6905.2  on 4999  degrees of freedom
## Residual deviance: 6848.6  on 4995  degrees of freedom
## AIC: 6858.6
##
## Number of Fisher Scoring iterations: 4
```

Pass it, along with the dataframe, to `additive_interactions`:

```
additive_interactions(m, dat=d, recode=TRUE)
```

```
## Warning in additive_interactions(m, dat = d, recode = TRUE): Recoding
## exposures; new reference category for E1 is 1 and for E2 is 1
```

```
##     Stat        Est       CI.lo      CI.hi      p.val.0 p.val.epi
## 1   RERI -0.2327401 -0.8126261 0.3471460 7.842540e-01         1
## 2     AP -0.1172736 -0.4163216 0.1817744 2.210617e-01        NA
## 3     E1  0.7691148  0.2367140 1.3015156 2.317267e-03        NA
## 4     E2  0.4672678  0.2741135 0.6604221 1.061065e-06        NA
## 5  E1:E2 -0.2363826 -0.8565017 0.3837365 7.725035e-01        NA
##   p.val.suff.cause
## 1        0.9999845
## 2               NA
## 3               NA
## 4               NA
## 5               NA
```

Note that interpretation of these results after recoding requires care; see Knol et al. (2011) for a detailed example.

References

1. Knol, M. J., VanderWeele, T. J., Groenwold, R. H., Klungel, O. H., Rovers, M. M., & Grobbee, D. E. (2011). Estimating measures of interaction on an additive scale for preventive exposures. *European Journal of Epidemiology*, 26(6), 433-438.