MANOVA Example

Below I compare ratings of three HMO (health maintenance organizations) using three measures of satisfaction: ease of choosing a personal physician (`ease`), recommend the health plan to others (`recom`), days waiting for routine care (`wait`). (From the welfare and health care study conducted by Karen Seccombe).[1]

Notice that Roy's largest root, which Olson (1976) warns will produce Type I errors, is the only significant effect.

```
glm ease recom wait by hmo
  /method=sstype(3)
 /intercept=include
  /print=descriptive etasq
 /design=hmo.
```

Between-Subjects Factors

		Value Label	N
hmo Health Maintenance Organization	1.00	HMO A	100
	2.00	HMO B	100
	3.00	HMO C	91

Descriptive Statistics

	hmo Health Maintenance Organization	Mean	Std. Deviation	N
ease Ease of choosing personal physician	1.00 HMO A	2.76000	1.264272	100
	2.00 HMO B	2.62000	1.170168	100
	3.00 HMO C	2.43956	.968490	91
	Total	2.61168	1.149393	291
recom Recommend health plan	1.00 HMO A	3.13000	1.050781	100
	2.00 HMO B	3.35000	.833333	100
	3.00 HMO C	3.45055	.806415	91
	Total	3.30584	.913085	291
wait Days waiting for Routine Care	1.00 HMO A	2.50000	1.210226	100
	2.00 HMO B	2.70000	1.321921	100
	3.00 HMO C	2.69231	1.546985	91
	Total	2.62887	1.359408	291

Multivariate Tests [a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.970	3034.526 [b]	3.000	286.000	<.001	.970
	Wilks' Lambda	.030	3034.526 [b]	3.000	286.000	<.001	.970
	Hotelling's Trace	31.831	3034.526 [b]	3.000	286.000	<.001	.970
	Roy's Largest Root	31.831	3034.526 [b]	3.000	286.000	<.001	.970
hmo	Pillai's Trace	.035	1.724	6.000	574.000	.113	.018
	Wilks' Lambda	.965	1.731 [b]	6.000	572.000	.111	.018
	Hotelling's Trace	.037	1.738	6.000	570.000	.110	.018
	Roy's Largest Root	.035	3.357 [c]	3.000	287.000	.019	.034

a. Design: Intercept + hmo
b. Exact statistic
c. The statistic is an upper bound on F that yields a lower bound on the significance level.

[1] Seccombe, K., Newsom, J.T., & Hoffman, K. (2006). Access to healthcare after welfare reform. *Inquiry, 43*, 167-179.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	ease Ease of choosing personal physician	4.903 [a]	2	2.451	1.867	.157	.013
	recom Recommend health plan	5.193 [b]	2	2.596	3.160	.044	.021
	wait Days waiting for Routine Care	2.533 [c]	2	1.266	.684	.506	.005
Intercept	ease Ease of choosing personal physician	1973.136	1	1973.136	1502.477	<.001	.839
	recom Recommend health plan	3182.283	1	3182.283	3873.821	<.001	.931
	wait Days waiting for Routine Care	2010.020	1	2010.020	1085.306	<.001	.790
hmo	ease Ease of choosing personal physician	4.903	2	2.451	1.867	.157	.013
	recom Recommend health plan	5.193	2	2.596	3.160	.044	.021
	wait Days waiting for Routine Care	2.533	2	1.266	.684	.506	.005
Error	ease Ease of choosing personal physician	378.218	288	1.313			
	recom Recommend health plan	236.587	288	.821			
	wait Days waiting for Routine Care	533.385	288	1.852			
Total	ease Ease of choosing personal physician	2368.000	291				
	recom Recommend health plan	3422.000	291				
	wait Days waiting for Routine Care	2547.000	291				
Corrected Total	ease Ease of choosing personal physician	383.120	290				
	recom Recommend health plan	241.780	290				
	wait Days waiting for Routine Care	535.918	290				

a. R Squared = .013 (Adjusted R Squared = .006)
b. R Squared = .021 (Adjusted R Squared = .015)
c. R Squared = .005 (Adjusted R Squared = -.002)

R

```
#super critical!!  make sure IV is a factor or results will be incorrect in car
> d$hmo <-as.factor(d$hmo)

#may need to install the tibble package separately to get Manova to work (car is supposed to include this)
library(tibble)
library(car)
model2 <- lm(cbind(ease, recom, wait) ~ hmo, data=d)
mod2sum <- Manova(model2, type="III")
summary(mod2sum, multivariate=TRUE)

Type III MANOVA Tests:

Sum of squares and products for error:
           ease     recom      wait
ease  378.21758 109.60198 -97.90769
recom 109.60198 236.58747 -48.38462
wait  -97.90769 -48.38462 533.38462

------------------------------------------

Term: (Intercept)

Sum of squares and products for the hypothesis:
         ease   recom  wait
ease  1049.76 1014.12 810.0
recom 1014.12  979.69 782.5
wait   810.00  782.50 625.0
```

```
Multivariate Tests: (Intercept)
                 Df test stat approx F num Df den Df                Pr(>F)
Pillai            1  0.883904 725.8266      3    286 < 0.000000000000000222
Wilks             1  0.116096 725.8266      3    286 < 0.000000000000000222
Hotelling-Lawley  1  7.613565 725.8266      3    286 < 0.000000000000000222
Roy               1  7.613565 725.8266      3    286 < 0.000000000000000222
------------------------------------------
 Term: hmo

Sum of squares and products for the hypothesis:
          ease    recom     wait
ease  4.902692 4.837885 2.845837
recom 4.837885 5.192596 3.415543
wait  2.845837 3.415543 2.532910

Multivariate Tests: hmo
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            2 0.0354030 1.723962      6    574 0.113063
Wilks             2 0.9646479 1.731155      6    572 0.111439
Hotelling-Lawley  2 0.0365949 1.738258      6    570 0.109856
Roy               2 0.0350906 3.356999      3    287 0.019298
```

Note: for factorial designs, the `manova` function does not use SS Type III, so use the `lm` function in the car package. The same model as above can be specified by `model2 <- lm(cbind(ease, recom, wait) ~ hmo, data=d)`. For factorial designs, a model specified on `lm`, can include multiple independent variables on the right side of the `~`. For example, if Factor A and Factor B, use `~A*B` on the right hand side instead of `hmo`.

Univariate tests need to be obtained separately, with `Anova()` in `car` or `ANOVA` in lessR or another method.

```
> mease = lm(ease ~ hmo, data = d)
> Anova(mease, type="III")
Anova Table (Type III tests)

Response: ease
             Sum Sq  Df  F value               Pr(>F)
(Intercept) 1049.76   1 799.3570 <0.0000000000000002 ***
hmo            4.90   2   1.8666               0.1565
Residuals    378.22 288
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> mrecom = lm(recom ~ hmo, data = d)
> Anova(mrecom, type="III")
Anova Table (Type III tests)

Response: recom
            Sum Sq  Df   F value                Pr(>F)
(Intercept) 979.69   1 1192.5852 < 0.0000000000000002 ***
hmo           5.19   2    3.1605               0.04388 *
Residuals   236.59 288
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> mwait = lm(ease ~ hmo, data = d)
> Anova(mwait, type="III")
Anova Table (Type III tests)

Response: ease
             Sum Sq  Df  F value               Pr(>F)
(Intercept) 1049.76   1 799.3570 <0.0000000000000002 ***
hmo            4.90   2   1.8666               0.1565
Residuals    378.22 288
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Example Write-up

A multivariate analysis of variance (MANOVA) was used to compare three health maintenance organizations on three dependent measures, ease of choosing a personal physician, recommending the health plan to others, and fewer days waiting for routine care. HMO C was generally more highly rated on each of the outcomes (M_1 = 3.24, M_2 = 3.13, M_3 = 2.50), followed by HMO B (M_1 = 3.38, M_2 = 3.35, M_3 = 2.70), and HMO A (M_1 = 3.56, M_2 = 3.45, M_3 = 2.69), for ease of choosing a physician, recommending the plan, and less waiting for care, respectively.[2] The multivariate test of the differences among the three groups was not significant, Pillai's trace = .035, $F(6,574) = 1.724$, $p = .11$, partial $\eta^2 = .018$. Of the univariate tests, only recommending the plan was significant, $F(2,288) = 3.16$, $p = .04$, partial $\eta^2 = .02$, accounting for approximately 2% of the variance in the multivariate outcomes. The univariate tests for ease of choosing a physician, $F(2,288) = 1.873$, $p = .16$, partial $\eta^2 = .01$, and fewer days waiting, $F(2,288) = .68$, $p = .51$, partial $\eta^2 = .01$, were nonsignificant.

References

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin, 83*, 579-586.

[2] You might use a Table to present the mean values instead, particularly for more dependent variables and groups.