Testing Simple Effects in ANOVA

Graduate Statistics

February 11, 2016

Tooth Growth and Vitamin C in Guinea Pigs

For this example, we'll be using the 2×3 factorial experiment described in the `ToothGrowth` dataset. We're interested in the how the following variables (and their interaction) affect tooth growth (`len`) in guinea pigs:

- `dose` – vitamin c dose in milligrams with three levels: 0.5 mg, 1 mg, or 2 mg.
- `supp` – form of vitamin c supplementation with two levels: orange juice ("`OJ`") or ascorbic acid ("`VA`").

In order to perform our analyses, we create the following set of orthogonal contrast codes:

	Orange Juice (OJ)			Ascorbic Acid (VC)		
	0.5 mg	1.0 mg	2.0 mg	0.5 mg	1.0 mg	2.0 mg
C_{sup}	+1	+1	+1	-1	-1	-1
C_{lin}	-1	0	+1	-1	0	+1
C_{quad}	-1	+2	-1	-1	+2	-1
$C_{s \times l}$	-1	0	+1	+1	0	-1
$C_{s \times q}$	-1	+2	-1	+1	-2	+1

However, we also want to test the *simple effect* of supplementation form for guinea pigs in the .05 mg dose condition. To address this, we construct the following set of dummy codes:

	Orange Juice			Ascorbic Acid		
	0.5 mg	1.0 mg	2.0 mg	0.5 mg	1.0 mg	2.0 mg
D_{sup}	1	1	1	0	0	0
D_{1mg}	0	1	0	0	1	0
D_{2mg}	0	0	1	0	0	1
$D_{s \times 1}$	0	1	0	0	0	0
$D_{s \times 2}$	0	0	1	0	0	0

What precisely are we interested in when we examine a simple effect like this? One way to think about it as the difference between the values we'd predict for animals in the 0.5 mg OJ and the .05 VC conditions. That is,

$$\widehat{y}_{\substack{dose=.5 \\ sup=OJ}} - \widehat{y}_{\substack{dose=.5 \\ sup=VC}} = ?$$

First, we'll try to examine this effect using our contrast codes. Note that an 0.5 mg OJ animal would have the green values on our contrast codes whereas an 0.5 mg VC would have the blue values:

	Orange Juice (OJ)			Ascorbic Acid (VC)		
	0.5 mg	1.0 mg	2.0 mg	0.5 mg	1.0 mg	2.0 mg
C_{sup}	+1	+1	+1	-1	-1	-1
C_{lin}	-1	0	+1	-1	0	+1
C_{quad}	-1	+2	-1	-1	+2	-1
$C_{s \times l}$	-1	0	+1	+1	0	-1
$C_{s \times q}$	-1	+2	-1	+1	-2	+1

Thus, after estimating our full model, we'd obtain the following predicted values:

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=OJ}} &= b_0 + b_1 C_{sup} + b_2 C_{lin} + b_3 C_{quad} + b_4 C_{s\times l} + b_5 C_{s\times q}\\ &= b_0 + b_1(1) + b_2(-1) + b_3(-1) + b_4(-1) + b_5(-1)\\ &= b_0 + b_1 - b_2 - b_3 - b_4 - b_5\end{aligned}$$

and

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=VC}} &= b_0 + b_1 C_{sup} + b_2 C_{lin} + b_3 C_{quad} + b_4 C_{s\times l} + b_5 C_{s\times q}\\ &= b_0 + b_1(-1) + b_2(-1) + b_3(-1) + b_4(1) + b_5(1)\\ &= b_0 + b_1 - b_2 - b_3 + b_4 + b_5\end{aligned}$$

Thus,

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=OJ}} - \widehat{y}_{\substack{dose=.5\\sup=VC}} &= b_0 + b_1 - b_2 - b_3 - b_4 - b_5 - [b_0 - b_1 - b_2 - b_3 + b_4 + b_5]\\ &= 2b_1 - 2b_4 - 2b_5\end{aligned}$$

Hmmm. Our predicted difference is $-2(b_4 + b_5)$. It isn't clear where our regression output would test this difference. What happens when we repeat the same procedures with our dummy codes?

	Orange Juice			Ascorbic Acid		
	0.5 mg	1.0 mg	2.0 mg	0.5 mg	1.0 mg	2.0 mg
D_{sup}	1	1	1	0	0	0
D_{1mg}	0	1	0	0	1	0
D_{2mg}	0	0	1	0	0	1
$D_{s\times 1}$	0	1	0	0	0	0
$D_{s\times 2}$	0	0	1	0	0	0

We then have

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=OJ}} &= b_0 + b_1 D_{sup} + b_2 D_{1mg} + b_3 D_{2mg} + b_4 D_{s\times 1} + b_5 D_{s\times 2}\\ &= b_0 + b_1(1) + b_2(0) + b_3(0) + b_4(0) + b_5(0)\\ &= b_0 + b_1\end{aligned}$$

and

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=VC}} &= b_0 + b_1 D_{sup} + b_2 D_{1mg} + b_3 D_{2mg} + b_4 D_{s\times 1} + b_5 D_{s\times 2}\\ &= b_0 + b_1(0) + b_2(0) + b_3(0) + b_4(0) + b_5(0)\\ &= b_0\end{aligned}$$

Now,

$$\begin{aligned}\widehat{y}_{\substack{dose=.5\\sup=OJ}} - \widehat{y}_{\substack{dose=.5\\sup=VC}} &= b_0 + b_1 - b_0\\ &= b_1\end{aligned}$$

Now, b_1 is directly represents the difference between the two cells we're interested in!

Note: this document was created with `knitr` and LyX.