7.8 Mathematical Complements

7.8.1 Iterated Projections

The Tower Property of conditional expectations states that

$$E\left[E(V|U_1,U_2) \mid U_1\right] = E(V \mid U_1) \tag{7.47}$$

This can be shown quite simply and elegantly using the vector space methods of Sections 1.19.5.5 and 2.12.8, as follows. (Readers may wish to review Section 2.12.8 before continuing.) Our only assumption is that the variables involved have finite variances, so that the vector space in 2.12.8 exists.

Let $\mathcal{A}$ and $\mathcal{B}$ denote the subspaces spanned by all functions of (U_1, U_2) and U_1, respectively. Denote the full space as $\mathcal{C}$.

Write

$$u = V \tag{7.48}$$

$$v = E(V|U_1,U_2) \tag{7.49}$$

$$w = E\left[E(V|U_1,U_2) \mid U_1\right] \tag{7.50}$$

Note that $\mathcal{B}$ is a subspace of $\mathcal{A}$. Indeed, w, a vector in $\mathcal{B}$, is the projection of v, a vector in $\mathcal{A}$, onto $\mathcal{B}$.

What we need to show is that w is also the projection of u onto $\mathcal{B}$. Since projections are unique, we will accomplish this if we show that

$$(w, u-w) = 0 \tag{7.51}$$

Write

$$u - w = (u-v) + (v-w) \tag{7.52}$$

Now consider each of the two terms on the right, which we will show are orthogonal to w. First,

$$(w, v - w) = 0 \tag{7.53}$$

because w is the projection of v onto $\mathcal{B}$. Second, since v is the projection of u onto $\mathcal{A}$, we know that

$$(q, u - v) = 0 \tag{7.54}$$

for any $q \epsilon \mathcal{A}$, including the case $q = w$. Thus we have

$$(w, u - w) = 0 \tag{7.55}$$

as desired, completing the proof.

7.8.2 Standard Errors for RFA

Recall that RFA is defined as

$$\widehat{\nu} = \frac{1}{r} \sum_{i=1}^{r} \widehat{\mu}(Q_i) \tag{7.56}$$

We will assume that $\mu(t)$ is linear in t, with the notation of Section 2.4.2. And for convenience, assume a model with no intercept term (Section 2.4.5). Then (7.56) becomes

$$\widehat{\nu} = \frac{1}{r} \sum_{i=1}^{r} Q_i' \widehat{\beta} = \overline{Q}' \widehat{\beta} \tag{7.57}$$

where $\overline{Q}$ is the (vector-valued) sample mean of the Q_i.

As in Section 2.8, our standard errors will be conditional on the X_j. We will also assume that the Q_i are i.i.d. and independent of the (Y_i, X_i) from which $\widehat{\beta}$ is computed.