

# Powell's "Projection" Theorem for a class of U-Statistics

**Theorem 1** (Lemma A.3 in Ahn and Powell (1993), see also Lemma 3.1 in Powell, Stock and Stoker (1989)) *For an iid sample  $\{X_i\}_{i=1}^n$  of random variables, consider an  $m^{\text{th}}$ -order U-statistic of the form*

$$U_n = \binom{n}{m}^{-1} \sum_c p_n(X_{i_1}, X_{i_2}, \dots, X_{i_m})$$

where the sum is taken over the set "c" of all  $\binom{n}{m}$  combinations of  $m$  distinct elements  $\{i_1, i_2, \dots, i_m\}$  from the set  $\{1, 2, \dots, n\}$ . Without loss of generality, the function  $p_n(\cdot)$  can be taken to be symmetric in its  $m$  arguments. Define the "projection" of this U-statistic as:

$$\widehat{U}_n \equiv \theta_n + \frac{m}{n} \sum_{i=1}^n [r_n(X_i) - \theta_n],$$

where

$$r_n(X_i) = E[p_n(X_i, X_{i_2}, \dots, X_{i_m}) | X_i], \quad \theta_n = E[r_n(X_i)]$$

Note that  $r_n(X_i)$  is the conditional expectation of  $p_n(\cdot)$  with respect to its first-argument, which is irrelevant because  $p_n(\cdot)$  is assumed symmetric: If  $m = 2$  for example,  $p_n(X_i, X_j) = p_n(X_j, X_i)$  and therefore  $r_n(X_i) = E[p_n(X_i, X_j) | X_i] = E[p_n(X_j, X_i) | X_i]$ .

If  $p_n(\cdot)$  satisfies:

$$E[\|p_n(X_1, \dots, X_m)\|^2] = O(n),$$

then

$$(i) \quad U_n = \theta_n + o_p(1)$$

$$(ii) \quad U_n = \widehat{U}_n + o_p(n^{-1/2})$$