1 Recap

Define the following:

\[ h_c(x_1, \ldots, x_c) = E(h(x_1, \ldots, x_c, X_{c+1}, \ldots, X_r)) \]

\[ \zeta_c = Var(h_c(X_1, \ldots, X_c)) \]

Now consider a U-Statistic:

\[ U_n = \frac{1}{n^r} \sum_{\beta} h(X_{\beta_1}, \ldots, X_{\beta_r}) \]

where \( E(h) = \theta \) and

\[ \text{Var}(U_n) = \left( \frac{n}{r} \right)^{-2} \sum_{c=0}^{r} \binom{n}{c} \binom{r}{c} (n-r) \zeta_c \]

Note that

\[ \text{Var}(U_n) = \frac{r^2 \zeta_1}{n} + O(n^{-2}) \]

1.1 Rao-Blackwellization

Note that we can write \( U_n = E(h(X_1, \ldots, X_r)|X_1, \ldots, X_r) \). Thus, we have the following inequality:

\[ E(U_n^2) = E(Eh(X_1, \ldots, X_r)|X_1, \ldots, X_r)^2 \leq E(Eh^2(X_1, \ldots, X_r)|X_1, \ldots, X_r) = h^2 \]

2 Projections

Define \( \mathcal{L}_2(P) \) as the set of functions that are finite when squared, and let \( T \) and \( \{S : S \in \mathcal{S}\} \) belong to \( \mathcal{L}_2(P) \).

**Definition 1.** \( \hat{S} \in \mathcal{S} \) is a projection of \( T \) on \( \mathcal{S} \) if and only if \( E((T - \hat{S})S) = 0 \) for all \( S \in \mathcal{S} \)

**Corollary 2 (From van der Vaart Chapter 11).** \( E(T^2) = E(T - \hat{S})^2 + E(\hat{S}^2) \)
Now consider a sequence of statistics $T_n$ and spaces $\mathcal{S}_n$ (that contain constant real variables) with projections $\hat{S}_n$.

**Theorem 3.** If $\frac{\text{Var}(T_n)}{\text{Var}(\hat{S}_n)} \to 1$ then

$$\frac{T_n - E(T_n)}{\text{std}(T_n)} - \frac{\hat{S}_n - E(\hat{S}_n)}{\text{std}(\hat{S}_n)} \overset{P}{\to} 0$$

**Proof:** Let $A_n = \frac{T_n - E(T_n)}{\text{std}(T_n)} - \frac{\hat{S}_n - E(\hat{S}_n)}{\text{std}(\hat{S}_n)}$. Note that $E(A_n) = 0$ and

$$\text{Var}(A_n) = 2 - 2 \left( \frac{\text{Cov}(T_n, \hat{S}_n)}{\text{std}(T_n)\text{std}(\hat{S}_n)} \right)$$

Since $(T_n - \hat{S}_n) \perp \hat{S}_n ((T_n - \hat{S}_n)$ is orthogonal to $\hat{S}_n)$, we have:

$$E(T_n \hat{S}_n) = E(\hat{S}_n^2) \Rightarrow$$

$$\text{Cov}(T_n, \hat{S}_n) = \text{Var}(\hat{S}_n) \Rightarrow$$

$$A_n \overset{p}{\to} 0$$

### 2.1 Conditional Expectations are Projections

$\mathcal{S}$ is the linear space of all measurable functions $g(Y)$ of $Y$. Define $E(X|Y)$ as a measurable function of $Y$ that satisfies $E(X - E(X|Y))g(Y) = 0$. As a consequence, we have the following:

- Setting $g \equiv 1$, then $E(X - E(X|Y)) = 0 \Rightarrow E(X) = E(E(X|Y))$

- $E(f(Y)X|Y) = f(Y)E(X|Y)$ because $E[f(Y)X - f(Y)E(X|Y)]g(Y) = E(X - E(X|Y))f(Y)g(Y) = 0$

- $E(E(X|Y,Z)|Y) = E(X|Y)$

### 2.2 Hájek Projections

Let $X_1, X_2, \ldots, X_n$ be independent, $\mathcal{S} = \{\sum_{i=1}^n g_i(x_i) : g_i \in \mathcal{L}_2(P)\}$. $\mathcal{S}$ is a Hilbert space.

**Lemma 3 (11.10 in van der Vaart).** Let $T$ have a finite 2nd moment. Then

$$\hat{S} = \sum_{i=1}^n E(T|X_i) - (n-1)E(T)$$

**Proof:**

$$E(E(T|X_i)|X_j) = \begin{cases} E[E(T|X_i)] = E(T) & \text{if } i \neq j \\ E(T|X_i) & \text{if } i = j \end{cases}$$

$$E(\hat{S}|X_j) = \sum_{i \neq j} E(T) - (n-1)E(T) + E(T|X_j) = E(T|X_j)$$

Thus we have that

$$E[(T - \hat{S})g(X_j)] = E[(E(T - \hat{S})|X_j)g(X_j)] = 0.$$ 

And we conclude $(T - \hat{S}) \perp \mathcal{S}$. 
3 Asymptotic Normality of U-Statistics

Assume \( E(h^2) < \infty \). Take Hájek projection of \((U_n - \theta)\) onto \( \{\sum_{i=1}^n g_i(x_i) : g_i \in \mathcal{L}_2(P)\}\). Define \( \hat{U}_n = U_n - \theta = \sum_{i=1}^n E((U - \theta)|X_i) \). We have that

\[
E(h(X_{\beta_1}, \ldots, X_{\beta_r}) - \theta|X_i = x) = \begin{cases} 
  h_1(x) & \text{if } i \in \beta \\
  0 & \text{otherwise}
\end{cases}
\]

Where \( h_1(x) = E(h(x_1, X_2, \ldots, X_r) - \theta) \). Now

\[
E(U_n - \theta|X_i) = \frac{1}{\binom{n}{r}} \sum_{\beta} E(h(x_{\beta_1}, \ldots, x_{\beta_r}|X_i) - \theta) = \frac{\binom{n-1}{r-1}}{\binom{n}{r}} = \frac{r}{n} h_1(x_i) \Rightarrow 
\]

\[
\hat{U}_n = \frac{r}{n} \sum_{i=1}^n h_1(x_i) 
\]

Note that \( E\hat{U}_n = 0 \) and

\[
\text{Var}(\hat{U}_n) = \frac{r^2}{n^2} [n[\text{Var}(h(X_1))]] = \frac{r^2}{n} \zeta_1 
\]

And so we have \( \frac{\text{Var}(U_n)}{\text{Var}(\hat{U}_n)} \rightarrow 1 \). By our previous theorem we have that

\[
\frac{U_n - \theta}{(\frac{r^2}{n^2} \zeta_1 + O(n^{-2}))^{\frac{1}{2}}} - \frac{\hat{U}_n}{(\frac{r^2}{n^2} \zeta_1)^{\frac{1}{2}}} \xrightarrow{P} 0
\]

By Slutsky we have

\[
\sqrt{n}(U_n - \theta - \hat{U}_n) \xrightarrow{P} 0 
\]

By CLT we have

\[
\sqrt{n}\hat{U}_n \xrightarrow{d} N(0, r^2 \zeta_1) 
\]

And by Slutsky again we have

\[
\sqrt{n}(U_n - \theta) \xrightarrow{d} N(0, r^2 \zeta_1) 
\]