LECTURE XIV. A THIRD ABSTRACT NORMAL APPROXIMATION THEOREM

In order to see the relation of the results of the tenth lecture to the abstract formulation of the first lecture it will be necessary to introduce fairly elaborate formalism. Roughly speaking, in order to obtain an exchangeable pair as in the first lecture we introduce a new random point conditionally independent of the original one given $\mathcal{C}$ and with the same conditional distribution given $\mathcal{C}$. This requires a new sample space, big enough to carry the original σ-algebra $\tilde{\mathcal{B}}$ and its copy $\tilde{\mathcal{B}}'$ corresponding to the new random point. The resulting structure seems quite formidable, but I believe it will be useful in the long run, although the simpler treatment of the tenth lecture and the even simpler treatment introduced at the end of the first lecture should suffice for many problems.

Let $(\tilde{\tilde{\Omega}},\tilde{\tilde{\mathcal{B}}},\tilde{\tilde{P}})$ be a probability space, let $\tilde{\mathcal{B}}$, $\tilde{\mathcal{B}}'$, and $\mathcal{C}$ be sub-σ-algebras of $\tilde{\tilde{\mathcal{B}}}$ and suppose that, under $\tilde{\tilde{P}}$, $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{B}}'$ are conditionally independent given $\mathcal{C}$. Also let $\mathcal{B}$ be a sub-σ-algebra of $\tilde{\mathcal{B}}$ and $\mathcal{B}'$ a sub-σ-algebra of $\tilde{\mathcal{B}}'$, and let $\gamma: \tilde{\tilde{\Omega}} \to \tilde{\tilde{\Omega}}$ be an involution, that is

(1) $$\gamma^2 = I_{\tilde{\tilde{\Omega}}} ,$$

such that

(2) $$\gamma^{-1} B \in \tilde{\tilde{\mathcal{B}}} \quad \text{for all} \quad B \in \tilde{\tilde{\mathcal{B}}},$$

(3) $$B \in \tilde{\mathcal{B}} \Leftrightarrow \gamma^{-1} B \in \tilde{\mathcal{B}}',$$

(4) $$B \in \mathcal{B} \Leftrightarrow \gamma^{-1} B \in \mathcal{B}',$$

$$\tilde{\tilde{P}} \circ \gamma^{-1} = \tilde{\tilde{P}}, \tag{5}$$

and

$$\gamma^{-1} C = C \quad \text{for all} \quad C \in \mathcal{C}. \tag{6}$$

Thus γ is a $\tilde{\tilde{P}}$-measure preserving involution of the measurable space $(\tilde{\tilde{\Omega}},\tilde{\tilde{\mathcal{B}}})$ that interchanges $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{B}}'$ and also interchanges $\mathcal{B}$ and $\mathcal{B}'$ and leaves each $\mathcal{C}$-measurable set invariant. Also let G be a real-valued $\tilde{\mathcal{B}}$-measurable random variable with

$$E|G| < \infty, \tag{7}$$

and define

$$W = E^{\mathcal{B}}G. \tag{8}$$

For any $\tilde{\tilde{\mathcal{B}}}$-measurable random variable Z I shall write

$$Z' = Z \circ \gamma^{-1}. \tag{9}$$

The conditions formulated in this paragraph will be referred to as the basic assumption of this lecture.

Now I can develop a sequence of lemmas, a theorem, and a corollary that are completely analogous to those of the tenth lecture. Only Lemma 1 will be proved since the proofs of the other results are obtained from the corresponding proofs of earlier lectures by obvious modification.

<u>Lemma 1</u>: In addition to the second basic assumption suppose $f: R \to R$ is a bounded, Borel-measurable function. Then

$$EWf(W) = EG(f(W)-f(W')) + E(E^{\mathcal{C}}G)(E^{\mathcal{C}}f(W)). \tag{10}$$

Proof:

$$\begin{aligned} EWf(W) &= E(E^{\mathcal{B}}G)f(W) = EGf(W) \\ &= EG(f(W)-f(W')) + EGf(W') \\ &= EG(f(W)-f(W')) + E(E^{\mathcal{C}}G)(E^{\mathcal{C}}f(W)). \end{aligned} \tag{11}$$

The first equality uses (8) and the second uses the fact that, again by (8), W is $\mathcal{B}$-measurable. The third equality is trivial and the fourth uses the conditional independence given $\mathcal{C}$ of the $\tilde{\mathcal{B}}$-measurable random variable G and the $\tilde{\mathcal{B}}'$-measurable random variable W'.

Lemma 2: In addition to the second basic assumption, suppose $h: R \to R$ is a Borel-measurable function such that, for some positive constant C we have

$$(12) \qquad |h(w)| \leq C(1+|w|)$$

for all $w \in R$. Let N and U_N be defined as in (II.2) and (II.4) and for brevity let

$$(13) \qquad f = U_N h.$$

Then f is bounded and

$$(14) \qquad Eh(W) = Nh + E[f'(W)-G(f(W)-f(W'))] - E(E^{\mathcal{C}}G)(E^{\mathcal{C}}f(W)).$$

The proof is analogous to that of Lemma I.4.

Lemma 3: In addition to the second basic assumption suppose $h: R \to R$ is absolutely continuous and satisfies (12) and also that

$$(15) \qquad E|G|(W-W')^2 < \infty.$$

Then, with f as in Lemma 2,

$$(16) \qquad Eh(W) = Nh - E(E^{\mathcal{C}}G)(E^{\mathcal{C}}f(W)) + Ef'(W)[1-G(W-W')] + \int EG(z-W')[\mathcal{I}\{z \leq W\} - \mathcal{I}\{z \leq W'\}]f''(z)dz.$$

The proof is analogous to that of Lemma III.1.

Theorem 1: In addition to the second basic assumption suppose $h: R \to R$ is bounded and absolutely continuous. Then

$$(17) \qquad |Eh(W) - Nh| \leq \sqrt{\frac{\pi}{2}}\sup|h-Nh|E|E^{\mathcal{C}}G| + 2\sup|h-Nh|\sqrt{E[1-E^{\mathcal{B}}G(W-W')]^2} + \frac{1}{2}\sup|h'|E|G|(W-W')^2,$$

and, for all real w_0,

$$(18) \qquad |P\{W \le w_0\} - \Phi(w_0)| \le \sqrt{\frac{\pi}{2}}\, E|E^{\mathcal{C}}G| + 2\sqrt{E[1-E^{\mathcal{B}}G(W-W')]^2} + \sqrt{\frac{2}{\pi}}\sqrt{E|G|(W-W')^2}\,.$$

The proof is analogous to that of Theorem III.1.

Now I shall try to indicate how this can be thought of as a special case of the formalism of diagram (I.28). We start from the definition of $\alpha\colon \mathfrak{F}_0 \to \mathfrak{F}$ by

$$(19) \qquad (\alpha f)(\omega) = \frac{1}{2}[G(\omega)-G'(\omega)][f(W(\omega))+f(W'(\omega))].$$

Clearly αf is antisymmetric in the sense that

$$(20) \qquad (\alpha f)\circ\gamma = -\alpha f.$$

Then we have

$$
\begin{aligned}
(21) \qquad 0 &= E\,\frac{1}{2}(G-G')[f(W)+f(W')] \\
&= E\,\frac{1}{2}(G-G')[2f(W)-(f(W)-f(W'))] \\
&= E[(G-G')f(W) - \frac{1}{2}(G-G')(f(W)-f(W'))] \\
&= EE^{\mathcal{B}}[(G-G')f(W) - \frac{1}{2}(G-G')(f(W)-f(W'))] \\
&= E[Wf(W)-(E^{\mathcal{C}}G)f(W) - \frac{1}{2}E^{\mathcal{B}}(G-G')(f(W)-f(W'))] \\
&= E[Wf(W)-(E^{\mathcal{C}}G)f(W) - E^{\mathcal{B}}G(f(W)-f(W'))].
\end{aligned}
$$

Now let us look at the separate steps of (21) and later at the way (21) fits into the formalism of diagram (I.28). On the second and third lines I have used trivial identities to rewrite the expression under the expectation sign. The crucial step is the introduction of the conditional expectation operator $E^{\mathcal{B}}$ on the fourth line, leading to the identity

$$(22) \qquad EE^{\mathcal{B}}[(G-G')f(W) - \frac{1}{2}(G-G')(f(W)-f(W'))] = 0.$$

In going from this to the next line I have evaluated the conditional expectation of the first of the two terms in brackets. First,

(23) $$E^{\mathcal{B}}Gf(W) = (E^{\mathcal{B}}G)f(W) = Wf(W),$$

since W is related to G by (8), and therefore is $\mathcal{B}$-measurable. Second,

(24) $$E^{\mathcal{B}}G'f(W) = (E^{\mathcal{B}}G')f(W) = (E^{\mathcal{C}}G')f(W)$$

since $\mathcal{B}$ and $\tilde{\mathcal{B}}'$ are conditionally independent given $\mathcal{C}$, and G' is $\tilde{\mathcal{B}}'$-measurable. Thus we obtain

(25) $$E[Wf(W)-(E^{\mathcal{C}}G)f(W) - \frac{1}{2} E^{\mathcal{B}}(G-G')(f(W)-f(W'))] = 0.$$

The last line of (21) is obtained by using the exchangeability of (G,W) with (G',W'):

(26) $$\begin{aligned} &E \frac{1}{2} E^{\mathcal{B}}(G-G')(f(W)-f(W')) \\ &= E \frac{1}{2} (G-G')(f(W)-f(W')) \\ &= EG(f(W)-f(W')) = EE^{\mathcal{B}}G(f(W)-f(W')). \end{aligned}$$

Now let us see how (21)-(26) fit into the formalism of diagram (I.28). The identity (22) fits the top line of this diagram exactly, asserting that

(27) $$E \circ T = 0,$$

with $T: \mathfrak{F} \to \mathfrak{X}$ defined, apart from the difficulty of matching probabilistic and algebraic notation, by

(28) $$T = E^{\mathcal{B}}.$$

Of course E is just the appropriate expectation mapping. Also (25) is obtained from (22) by straightforward evaluation of the conditional expectations occurring in (22) so it too can be considered a realization of diagram (I.28). The final form of (21), which is essentially the same as (10) in Lemma 1 was included mainly in order to obtain a development closer to that of the tenth lecture. I believe (25) will eventually prove to be more useful.

The basic idea of diagram (I.28) may become clearer if it is given an even more abstract formulation. Let us look at the diagram

(29)
$$\begin{tikzcd}
\mathfrak{F} \arrow[r, "T"] & \mathcal{X} \arrow[r, "E"] & \mathcal{Y} \\
\mathfrak{F}_0 \arrow[u, "\alpha"] \arrow[r, shift left, "T_0"] & \mathcal{X}_0 \arrow[u, "\beta"] \arrow[l, shift left, "U_0"] \arrow[r, shift left, "E_0"] & \mathcal{Y}_0 \arrow[u, "\gamma"] \arrow[l, shift left, "\delta_0"]
\end{tikzcd}$$

where $\mathfrak{F}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathfrak{F}_0$, $\mathcal{X}_0$, and $\mathcal{Y}_0$ are linear spaces and all the arrows represent linear mappings. It is assumed that

(30)
$$E \circ T = 0$$

and

(31)
$$I_{\mathcal{X}_0} = T_0 \circ U_0 + \delta_0 \circ E_0$$

and

(32)
$$\gamma = E \circ \beta \circ \iota_0 .$$

Then

(33)
$$\begin{aligned}
0 &= E \circ T \circ \alpha \circ U_0 \\
&= E \circ \beta \circ T_0 \circ U_0 + E \circ (T \circ \alpha - \beta \circ T_0) \circ U_0 \\
&= E \circ \beta \circ (I_{\mathcal{X}_0} - \delta_0 \circ E_0) + E \circ (T \circ \alpha - \beta \circ T_0) \circ U_0 \\
&= (E \circ \beta - \gamma \circ E_0) + E \circ (T \circ \alpha - \beta \circ T_0) \circ U_0 .
\end{aligned}$$

Thus we have expressed $E \circ \beta - \gamma \circ E_0$, a measure of the departure from commutativity of the right-hand square, in terms of $T \circ \alpha - \beta \circ T_0$, a measure of the departure from commutativity of the left-hand square.

We can imagine this as arising in the following way. We start with only part of the diagram:

(34)
$$\begin{tikzcd} \mathfrak{F} \arrow[r, "T"] & \mathcal{X} \arrow[r, "E"] & \mathcal{Y} \\ & \mathcal{X}_0 \arrow[u, "\beta"] & \end{tikzcd}$$

We are given the linear mapping E: $\mathcal{X} \to \mathcal{Y}$ and typically $\mathcal{X}_0$ is a linear subspace of $\mathcal{X}$ and β is the appropriate inclusion mapping. Our aim is to approximate $E\circ\beta$: $\mathcal{X}_0 \to \mathcal{Y}$, and $\mathfrak{F}$ and T, satisfying (30) have been constructed to further that aim. Then the linear spaces $\mathfrak{F}_0$ and $\mathcal{Y}_0$ and the linear mappings T_0: $\mathfrak{F}_0 \to \mathcal{X}_0$ and E_0: $\mathcal{X}_0 \to \mathcal{Y}_0$ are constructed as a partial approximation to the top line of the diagram, with the connections α and γ. As indicated earlier, when the structure is completed by specifying U_0: $\mathcal{X}_0 \to \mathfrak{F}_0$, and δ_0: $\mathcal{Y}_0 \to \mathcal{X}_0$ satisfying (31) and (32), we are able to derive (33), which may provide a useful expression for $E\circ\beta$.

Let us look at the conditions (31) and (32). In the special case of diagram (I.28), condition (32) holds because $\mathcal{Y} = \mathcal{Y}_0 = R$, $\gamma = I_R$, β and δ_0 are inclusion mappings, and for constant c, $Ec = E_0c = c$ when E and E_0 are expectation mappings. In all the applications I have ever considered for this formalism the condition

(35) $$E_0\circ T_0 = 0$$

has been satisfied, which seems natural because the lower line of diagram (29) is intended as a partial approximation to the upper line. When (35) holds, (31) implies that

(36) $$T_0 = T_0\circ U_0\circ T_0,$$

that is, U_0 is a pseudo-inverse of T_0. Actually in all the applications considered in this set of notes (Lectures I-XV), U_0 is a right inverse of T_0, that is

(37) $$I_{\mathfrak{F}_0} = T_0 \circ U_0.$$

However in the multidimensional case this does not hold.

The first part of the lecture has been described adequately in the introductory paragraph. In the second part of the lecture, starting with diagram (29), I have tried to give a slightly more abstract version of the basic idea of diagram (I.28). I do not know of any applications for the increased generality but the study of conditional expectation seems to be a possibility. It may also be suitable for application to linear functional equations outside of probability theory.