Introduction to the Mathematical and Statistical Foundations of Econometrics

Remaining corrections and improvements in the 2004 and 2007 editions

January 10, 2014

Page 9, line 2 from top: Replace the first line of the equation with

\[ P(\{k\}) = \binom{K}{k} \binom{N-K}{n-k} = \frac{K!(N-K)!}{k!(K-k)!(n-k)!(N-K-n+k)!} \frac{N!}{n!(N-n)!} \]

(the actual correction is the bold \( k! \))

Page 10, line 3 from below: Replace \([0.5-1/q,0.5+1/q]\) with \((0.5-1/q,0.5+1/q]\)

Page 11, line 2 from top: Replace \([0.5-1/q,0.5+1/q]\) with \((0.5-1/q,0.5+1/q]\)

Page 11, line 4 from top: Replace \([0.5-1/q,0.5+1/q]\) with \((0.5-1/q,0.5+1/q]\)

Page 13, line 7 from below: Replace “each” with “at least one”

Page 13, line 6 from below: Replace “they belong” with “it belongs”

Page 18, line 8 from top: Replace “probability” with “unique probability”

Page 18, line 10 from top: Delete this line

Page 21, line 5 of Def. 1.8: Replace \( \mathcal{F}^{k} \) with \( \mathcal{F}^{k} \).

Page 25, line 10 from top: Replace \( \mu([\infty,a)) \) with \( \mu((\infty,a]) \)

Page 31, exercise 16: Replace “Definition 1.1.2” with “Definition 1.1.1”

Page 36, lines 6 and 7: Replace these two lines with the following:

---

1 Page and line numbers refer to the 2007 edition. Thanks to Dongkoo Kim and Soo Hyun Oh for pointing out many errors.

2 It was stated that the proof of Lemma 1.B.4. is too difficult and too long. However, there is a much shorter proof in Jeffrey S. Rosenthal: A First Look at Rigorous Probability Theory (2000), World Scientific Publishing Company, p.14.

Thanks to David Jinkins for pointing this out to me.
Proof: Suppose that there exist two extensions of outer measures, $P_1$ and $P_2$, to probability measures on $\{\Omega, \mathcal{F}\}$ such that for all sets $B \in \mathcal{F}$, $P_1(B) = P_2(B) = P(B)$. Then by the definition of outer measure, for all sets $A \in \mathcal{F}$,

$$P_1(A) = \inf_{A \subseteq \bigcup_{i=1}^{n} B_i, B_i \in \mathcal{F}_0} \sum_{i=1}^{n} P(B_i) = \inf_{A \subseteq \bigcup_{i=1}^{n} B_i, B_i \in \mathcal{F}_0} \sum_{i=1}^{n} P_2(B_i) \geq \inf_{A \subseteq \bigcup_{i=1}^{n} B_i, B_i \in \mathcal{F}_0} P_2\left(\bigcup_{i=1}^{n} B_i\right) \geq \inf_{A \subseteq \bigcup_{i=1}^{n} B_i, B_i \in \mathcal{F}_0} P_2(A) = P_2(A)$$

and similarly, $P_2(A) \geq P_1(A)$. Thus, $P_2(A) = P_1(A)$ for all sets $A \in \mathcal{F}$. Q.E.D.

Page 41, lines 3-4: Replace $\leq$ with $<$ (4 times), or replace these lines with

$$(b) \quad \{x \in \mathbb{R}: f_1(x) \leq y\} = \bigcap_{n=1}^{\infty} \bigcup_{j=1}^{n} \{x \in \mathbb{R}: g_j(x) < y + n^{-1}\} \in \mathcal{B}.

(c) \quad \{x \in \mathbb{R}: h_1(x) \leq y\} = \bigcap_{m=1}^{\infty} \bigcap_{n=1}^{m} \bigcup_{j=n}^{m} \{x \in \mathbb{R}: g_j(x) < y + m^{-1}\} \in \mathcal{B}.

Page 41, line 16 from top: Replace $\left(\inf_{x \in B(j,m,n)} g(x_j)\right)$ with $\left(\inf_{x \in B(j,m,n)} g(x_j)\right)$

Page 42, line 2 of Th. 2.7: Replace “limit” with “pointwise limit”

Page 42, line 1 of section 2.3: Replace $\sum_{j=1}^{m}$ with $\sum_{j=0}^{m}$

Page 42, line 3 of section 2.3: Replace $\sum_{j=1}^{m}$ with $\sum_{j=0}^{m}$ twice

Page 43, line 1 of Def. 2.5: Replace $\{\mathbb{R}^k, \mathcal{F}_k\}$ with $\{\mathbb{R}^k, \mathcal{F}_k\}$

Page 45, equation (2.11): Replace $\leq$ with $<$

Page 46, line 2 from below: Replace $X(\Omega)$ with $X(\omega)$

Page 46, line 2 of Th. 2.12: Replace “with probability 1” with ”, i.e., $Y(\omega) \neq 0$ for all $\omega \in \Omega$,”

Page 47, line 2 of Def. 2.9: Replace “of P” with “to P”

Page 49: Replace the proof of Theorem 2.18 with the following:
Let \( g(x) \) be a nonnegative Borel measurable function on \( \mathbb{R} \), and let \( X \) be a random variable defined on the probability space \( \{\Omega, \mathcal{F}, P\} \), with induced probability measure \( \mu_{\mathcal{F}}(.) \) defined on the Borel sets in \( \mathbb{R} \). Let \( S_n(g) \) be the set of all nonnegative simple functions \( g_n(x) = \sum_{j=1}^{n} a_j I(x \in B_j) \) with \( a_j \leq \inf_{x \in B_j} g(x) \), so that \( g_n(x) \leq g(x) \) pointwise in \( x \). Without loss of generality we may assume that \( 0 = a_1 < a_2 < \ldots < a_n \) and \( \bigcup_{j=1}^{n} B_j = \mathbb{R} \). Then by definition,

\[
\int g(x)d\mu_{\mathcal{F}}(x) = \sup_{g_{*} \in S_n(g)} \left\{ \int g_{*}(x)d\mu_{\mathcal{F}}(x) \right\}.
\]

Let \( Y = g(X) \), and denote by \( S_n(Y) \) be the set of all nonnegative simple random variables \( Y_n(\omega) = \sum_{j=1}^{n} b_j I(\omega \in A_j) \) with \( b_j \leq \inf_{\omega \in A_j} Y(\omega) \), so that \( Y_n(\omega) \leq Y(\omega) \) pointwise in \( \omega \in \Omega \). Again, without loss of generality we may assume that \( 0 = b_1 < b_2 < \ldots < b_n \) and \( \bigcup_{j=1}^{n} A_j = \Omega \). Then by definition,

\[
\int g(X(\omega))dP(\omega) = \int Y(\omega)dP(\omega) = \sup_{Y_{*} \in S_n(Y)} \left\{ \int Y_{*}(\omega)dP(\omega) \right\}.
\]

Each simple function \( g_n(x) = \sum_{j=1}^{n} a_j I(x \in B_j) \in S_n(g) \) corresponds to a simple random variable

\[
Y_{n}(\omega) = g_{n}(X(\omega)) = \sum_{j=1}^{n} a_j I(X(\omega) \in B_j) = \sum_{j=1}^{n} a_j I(\omega \in C_j) \in S_n(Y)
\]

where \( C_j = \{\omega \in \Omega: X(\omega) \in B_j\} \in \mathcal{F} \), such that

\[
\int Y_{n}(\omega)dP(\omega) = \sum_{j=1}^{n} a_j P(C_j) = \sum_{j=1}^{n} a_j \mu_{\mathcal{F}}(B_j) = \int g_{n}(x)d\mu_{\mathcal{F}}(x).
\]

Hence,

\[
\int g_{n}(x)d\mu_{\mathcal{F}}(x) \leq \sup_{Y_{*} \in S_n(Y)} \left\{ \int Y_{*}(\omega)dP(\omega) \right\} = \int Y(\omega)dP(\omega)
\]

and thus,

\[
\int g(x)d\mu_{\mathcal{F}}(x) = \sup_{g_{*} \in S_n(g)} \left\{ \int g_{*}(x)d\mu_{\mathcal{F}}(x) \right\} \leq \int Y(\omega)dP(\omega).
\]

On the other hand, each simple random variable \( Y_n(\omega) = \sum_{j=1}^{n} b_j I(\omega \in A_j) \in S_n(Y) \) corresponds to a simple function \( g_{n}^{*}(x) = \sum_{j=1}^{n} b_j I(x \in B_j) \in S_n(g) \) such that

\[
g_{n}^{*}(X(\omega)) = \sum_{j=1}^{n} b_j I(X(\omega) \in B_j) \geq Y_n(\omega)
\]

To see this, let

\[
B_j = \{x \in \mathbb{R}: g(x) \in [b_j, b_{j+1})\} \quad \text{for} \quad j = 1, \ldots, n-1, \quad B_n = \{x \in \mathbb{R}: g(x) \in [b_n, \infty)\}.
\]

Then

\[
g_{n}^{*}(X(\omega)) = \sum_{j=1}^{n} b_j I(X(\omega) \in B_j) = \sum_{j=1}^{n} b_j I(b_j \leq g(X(\omega)) < b_{j+1}) + b_n I(g(X(\omega)) \geq b_n)
\]

\[
= \sum_{j=1}^{n-1} b_j I(b_j \leq Y(\omega) < b_{j+1}) + b_n I(Y(\omega) \geq b_n)
\]
\[ \sum_{j=1}^{n-1} b_j I(b_j \leq Y_n(\omega) < b_{j+1}) + b_n I(Y_n(\omega) \geq b_n) \]
\[ = \sum_{j=1}^{n} b_j I(Y_n(\omega) = b_j) = \sum_{j=1}^{n} b_j I(\omega \in A_j) = Y_n(\omega). \]

Hence,
\[ \int Y_n(\omega) dP(\omega) \leq \int g^*_n(X(\omega)) dP(\omega) = \int g^*_n(x) d\mu_x(x) \]
\[ \leq \sup_{g^*_n \in \mathbb{R}} \int g^*_n(x) d\mu_x(x) = \int g(x) d\mu_x(x) \]
and thus
\[ \int Y(\omega) dP(\omega) = \sup_{Y^* \in \mathbb{R}} \int Y_n(\omega) dP(\omega) \leq \int g(x) d\mu_x(x). \]

Consequently,
\[ \int g(X(\omega)) dP(\omega) = \int Y(\omega) dP(\omega) = \int g(x) d\mu_x(x) \quad (2.16) \]
for all nonnegative Borel measurable functions \( g(x) \), and therefore (2.16) holds also for all Borel measurable functions \( g(x) \). Q.E.D.

---

**Page 50, line 1 of Def. 2.13:** Replace m’s with \( m \)-th

**Page 50, line 2 of Def. 2.13:** Replace \( E[X - \mu_x] \) with \( E[(X - \mu_x)^m] \)

**Page 51, line 2 from top:** Replace “Function” with “function”

**Page 56, proof of Th. 2.21:** The proof of Theorem 2.21 can be shortened considerably as follows:\(^3\)

Let \( a \) be a continuity point of \( F(x) \) and \( G(y) \). Replace \( b \) in (2.34) with \( a + 1/n \), i.e., replace \( \varphi(x) \) with

\[ \varphi_n(x|a) = \begin{cases} 
0 & \text{if } x \geq a + 1/n, \\
1 & \text{if } x < a, \\
n(a - x) + 1 & \text{if } a \leq x < a + 1/n.
\end{cases} \]  
(2.34)

Note that \( \lim_{n \to \infty} \varphi_n(x|a) = I(x \leq a) \). It follows therefore from Theorem 2.17 that

---

\(^3\) Thanks to Bruno Sultanum Teixeira for suggesting this.
\[
\lim_{n \to \infty} E[\varphi_n(X|a)] = E[I(X \leq a)] = F(a) \quad \text{and} \quad \lim_{n \to \infty} E[\varphi_n(Y|a)] = E[I(Y \leq a)] = G(a).
\]

Because \( E[\varphi_n(X|a)] = E[\varphi_n(Y|a)] \), it follows that \( F(a) = G(a) \).

---

Page 57, line 2 from top: Replace \( E[\varphi(X)] \) with \( E[\varphi(Y)] \)

Page 57, line 15 from top: Replace \( P_j \) with \( p_j \)

Page 59, line 1 of exercise 17: Replace \( E(E) \) with \( E(X) \)

Page 74, Th. 3.5: Replace \( ^4 P[E(Y|\mathcal{F}) = E(Y)] = 1 \) with \( E[Y|\mathcal{F}] = E[Y] \)

Page 74, Proof of Th. 3.5: Replace the proof with the following:

Denote \( Z = E[Y|\mathcal{F}] \). It is left as an exercise to prove that \( Z = E[Y] \) a.s., along the same lines as the proofs of Theorems 3.2 and 3.4. Next, recall from Definition 3.1 that \( Z \) is measurable \( \mathcal{F} \), i.e., for any Borel set \( B \) the set \( \{ \omega \in \Omega : Z(\omega) \in B \} \) is contained in \( \mathcal{F} \). Hence, for any Borel set \( B \) we have either \( \{ \omega \in \Omega : Z(\omega) \in B \} = \Omega \) or \( \{ \omega \in \Omega : Z(\omega) \in B \} = \emptyset \). Now let \( B \) be the singleton \( \{E[Y]\} \). Then either \( \{ \omega \in \Omega : Z(\omega) = E[Y] \} = \Omega \) or \( \{ \omega \in \Omega : Z(\omega) = E[Y] \} = \emptyset \). The latter is excluded by \( Z = E[Y] \) a.s., hence \( Z(\omega) = E[Y] \) for all \( \omega \in \Omega \). In other words, \( Z = E[Y] \) holds exactly.

Page 75, line 8 from top: \( = \mathcal{F}_{X,Z} \) with \( \subset \mathcal{F}_{X,Z} \)

Page 76, line 4 from below: Replace "discrete" with "simple"

Page 77, line 4 of part (b): Insert the following sentence after "monotonic."

The sequence \( X_n(\omega) \) can be constructed similar the simple function \( g_n(x) \) in the proof of Theorem 2.6, with \( g(x) \) replaced by \( X(\omega) \).

Page 83, Exercise 10: Replace "Borel-measurable" with "continuous"

Page 89, line 9 from below: Replace the expression for \( \varphi_{NB(m,p)}(t) \) with

---

\(^4\) Thanks to Renxiang Dai, Tilburg University, for suggesting this.
\[ \Phi_{NB(m,p)}(t) = \left( \frac{p}{1 - (1 - p)e^{it}} \right)^m = \left( \frac{p}{1 - (1 - p)\cos(t) - i(1 - p)\sin(t)} \right)^m = \left( \frac{p(1 - (1 - p)\cos(t) + i(1 - p)\sin(t))}{p^2 + 2(1 - p)(1 - \cos(t))} \right)^m \]

Page 96, last line: Replace \( m(i.t) \) with \( m_{N(0,1)}(i.t) \)

Page 97, line 2 from top: Replace \( m'(i.t) \) with \( m'_{N(0,1)}(i.t) \)

Page 97, line 2 from top: Replace \( m''(i.t) \) with \( m''_{N(0,1)}(i.t) \)

Page 99, line 12 from top: Replace \( N(1,n/y) \) with \( N(0,n/y) \)

Page 101, line 9 from below: Replace the expression for \( \Phi_{U[a,b]}(t) \) with

\[ \Phi_{U[a,b]}(t) = \frac{\exp(i.b.t) - \exp(i.a.t)}{i.(b - a)t} = \frac{(\sin(b.t) - \sin(a.t)) - i(\cos(b.t) - \cos(a.t))}{(b - a)t} \]

Page 104, line 1: Replace APPENDICES with APPENDIXES \(^5\)

Page 115, line 3 from below: Replace \( \begin{pmatrix} 1 & B^\top \\ 0 & I_k \end{pmatrix} \) with \( \begin{pmatrix} 1 & B^\top \\ 0 & I_k \end{pmatrix} \)

Page 140, equation (6.1): Replace the left-hand side of the inequality with

\[ E\left[ \left( \frac{1}{n} \sum_{j=1}^{n} (Z_j - E(Z_j)) \right)^2 \right] \]

Page 141, equation (6.2): Replace \( \sum_{k=1}^{j-1} \) with \( \sum_{k=1}^{j} \) twice

\(^5\) According to the British-English spelling.
Page 141, line 2 below eq. (6.2): Replace $\Sigma_{k=1}^j$ with $\Sigma_{k=1}^i$.

Page 141, line 2 below eq. (6.2): The “easy” equality $\Sigma_{k=1}^j k\alpha_k = \Sigma_{k=1}^i \Sigma_{i-k}^j \alpha_i$ appears to be not that obvious. Therefore, here is the proof:

\[
\sum_{k=1}^j k\alpha_k = \alpha_1 + (\alpha_2 + \alpha_2) + (\alpha_3 + \alpha_3 + \alpha_3) + \ldots + (\alpha_j + \alpha_j + \alpha_j + \ldots + \alpha_j)
\]

\[
= (\alpha_1 + \alpha_2 + \alpha_3 + \ldots + \alpha_j) + (\alpha_2 + \alpha_3 + \ldots + \alpha_j) + (\alpha_3 + \ldots + \alpha_j) + \ldots + \alpha_j
\]

\[
= \sum_{i=1}^j \alpha_i + \sum_{i=2}^j \alpha_i + \sum_{i=3}^j \alpha_i + \ldots + \sum_{i=j}^j \alpha_i = \sum_{i=1}^j \sum_{i=k}^j \alpha_i
\]

Page 142, line 3 in Th. 6.3: Replace in c with at $x = c$.

Page 143, Proof of Th. 6.5: Replace the sentence “Again, without loss of generality we may assume that $P[X = 0] = 1$ and that $X_n$ is a nonnegative random variable.” with the following:

It will be shown first that $X_n \xrightarrow{p} X$ and $\lim_{M \to \infty} \sup_{n \geq 1} E[|X_n| I(|X_n| > M)] = 0$ imply that $|X_n - X|$ is uniformly integrable. To show this, we need to show that $\sup_{n \geq 1} E[|X_n|] < \infty$ and $E[|X|] < \infty$, as follows. Let $M > 0$. Then

\[
E[|X_n|] = E[|X_n| I(|X_n| \leq M)] + E[|X_n| I(|X_n| > M)] \leq M + \sup_{n \geq 1} E[|X_n| I(|X_n| > M)]
\]

Because $X_n$ is uniformly integrable, for an arbitrary $\varepsilon > 0$ we can choose an $M_\varepsilon$ such that the second term is less than $\varepsilon$. Consequently, \( \sup_{n \geq 1} E[|X_n|] \leq M_\varepsilon + \varepsilon < \infty. \)

To show that $E[|X|] < \infty$, choose $K > 0$ and $\varepsilon > 0$ arbitrary. Then

\[
E[|X| I(|X| \leq K)] = E[|X| I(|X_n - X| > \varepsilon) I(|X| \leq K)] + E[|X| I(|X_n - X| \leq \varepsilon) I(|X| \leq K)]
\]

\[
\leq K.P[|X_n - X| > \varepsilon] + E[|X_n - X| I(|X_n - X| \leq \varepsilon) I(|X| \leq K)] + E[|X_n| I(|X_n| - X| \leq \varepsilon) I(|X| \leq K)]
\]

\[
\leq K.P[|X_n - X| > \varepsilon] + \varepsilon + \sup_{m \geq 1} E[|X_m|]
\]

Letting $n \to \infty$ it follows from $X_n \xrightarrow{p} X$ that $E[|X| I(|X| \leq K)] \leq \varepsilon + \sup_{m \geq 1} E[|X_m|]$ and next, letting $K \to \infty$, it follows that $E[|X|] \leq \varepsilon + \sup_{n \geq 1} E[|X_n|]$. Because $\varepsilon$ was arbitrary, it follows
now that $E[|X|] \leq \sup_{n \geq 1} E[|X_n|] < \infty$.

To show that $|X_n - X|$ is uniformly integrable, observe that for arbitrary $M > 0$ and $K > 0$,

$$E[|X_n - X|, I(|X_n - X| > M)] = E[|X_n - X|, I(|X_n - X| > M), I(|X_n| \leq K), I(|X| \leq K)]$$

$$+ E[|X_n - X|, I(|X_n - X| > M), I(|X_n| > K), I(|X| \leq K)]$$

$$+ E[|X_n - X|, I(|X_n - X| > M), I(|X_n| \leq K), I(|X| > K)]$$

$$+ E[|X_n - X|, I(|X_n - X| > M), I(|X_n| > K), I(|X| > K)]$$

$$\leq 4.K.P[|X_n - X| > M] + 2\sup_{m \geq 1} E[|X_m|, I(|X_m| > K)] + 2.E[|X|, I(|X| > K)]$$

Hence,

$$\limsup_{n \to \infty} E[|X_n - X|, I(|X_n - X| > M)] \leq 2.\sup_{m \geq 1} E[|X_m|, I(|X_m| > K)] + 2.E[|X|, I(|X| > K)]$$

and thus, letting $K \to \infty$, it follows that $\lim_{n \to \infty} E[|X_n - X|, I(|X_n - X| > M)] = 0$.

Next, choose an arbitrary $\varepsilon > 0$, and pick an $M_0 > 0$. Then there exists a natural number $n_0(\varepsilon)$ such that $E[|X_n - X|, I(|X_n - X| > M_0)] < \varepsilon$ for all $n > n_0(\varepsilon)$, and therefore also $E[|X_n - X|, I(|X_n - X| > M)] < \varepsilon$ for all $n > n_0(\varepsilon)$ and $M > M_0$.

Hence,

$$\lim_{M \to \infty} \sup_{n \geq 1} E[|X_n - X|, I(|X_n - X| > M)] \leq \varepsilon + \max_{1 \leq n \leq n_0(\varepsilon)} \lim_{M \to \infty} E[|X_n - X|, I(|X_n - X| > M)]$$

$$= \varepsilon$$

where the equality follows from the fact that $E[|X_n - X|] < \infty$. Because $\varepsilon$ was arbitrary, it follows now that $|X_n - X|$ is uniformly integrable. Therefore, without loss of generality we may now assume that $P[X = 0] = 1$ and that $X_n$ is a nonnegative random variable.

**Page 144, last line of Th. 6.7:** Replace this line with

$P(X \in B) = 1$. Then $\Psi(X_n) \to \Psi(X)$ a.s.

**Page 146, first line below (6.8):** Replace “Theorem 6.3” with “Theorem 6.10”
The reason is that the singleton $B = \{c\}$ is closed and bounded, and that for this case Theorem 6.12 may not hold. If $B$ is open then there exists a $\delta > 0$ such that $\{x \in \mathbb{R}^k : \|x-c\| < \delta\} \subset B$, which is an essential element of the proof of Theorem 6.12.

---

Page 147, equation (6.11): Replace equation (6.11) with:

$$E[g(X_1, \theta)] = \int_{-\infty}^{\infty} \frac{\exp(-x^2/2\sqrt{2\pi})}{\sqrt{2\pi}} dx = \int f(x, \theta) h(x, \theta) dx = \gamma(\theta - \theta_0),$$

(6.11)

Page 149, line 1 from top: Replace the sentence “where $B$ is a closed and bounded subset of $\mathbb{R}^k$ containing $c$” with “where $B$ is an open subset of $\mathbb{R}^k$ containing $c$”.

Page 150, line 7 from top: Replace $X_n \rightarrow_p X$ with $X_n \rightarrow_p X$.

Page 150, Proof of Th. 6.17: Replace the proof of Theorem 6.17 with the following:

**Proof:** Let $X_n$ and $X$ be random vectors in $\mathbb{R}^k$. It follows from Theorem 6.B.3 in Appendix 6.B that for each subsequence $n_j$ of $n$ there exists a further subsequence $n_{jm}$ such that $X_{n_{jm}} \rightarrow X$ a.s. as $m \rightarrow \infty$, so that by Theorem 6.7,

$$\varphi\left(\left\{X_{n_{jm}}\right\}\right) \rightarrow \varphi(X) \text{ a.s. as } m \rightarrow \infty,$$

for every bounded and continuous function $\varphi(x)$ on $\mathbb{R}^k$. This implies, by Theorem 6.B.3, that $\varphi(X_n) \rightarrow_p \varphi(X)$, which by Theorem 6.4 implies that $\lim_{n \rightarrow \infty} E[\varphi(X_n)] = E[\varphi(X)]$. Theorem 6.17 now follows from Theorem 6.18. Q.E.D.

Page 153, line 9 from top: Replace $|E[\Phi(X_n, Y_n)] - E[\Phi(X_n, c)]|$ with $|E[\Phi(X_n, Y_n)] - E[\Phi(X_n, c)]|$.

Page 154, line 12 from top: Replace $\mathbb{R}^{k \times k}$ with $\mathbb{R}^{k^2}$.

Page 157, last line of Th. 6.25: Replace $\varphi(X_n)$ with $\Phi(X_n)$.

Page 159, line 7 from below: Replace “Theorem 6.1” with “Theorem 6.28”.

Page 161, last line: Replace “Theorem 6.21” with “Theorem 6.12”.

---

The reason is that the singleton $B = \{c\}$ is closed and bounded, and that for this case Theorem 6.12 may not hold. If $B$ is open then there exists a $\delta > 0$ such that $\{x \in \mathbb{R}^k : \|x-c\| < \delta\} \subset B$, which is an essential element of the proof of Theorem 6.12.
Page 180, line 7 from top: Replace “The exact proof” with “A more exact proof”

Page 180, line 10 from top: Replace “It is possible” with “Suppose it is possible”

Page 180, line 11 from top: Insert the following sentences after “minimal.”:

This is not always possible. See http://econ.la.psu.edu/~hbierens/WOLD.PDF for a counter example.

Page 182, lines 8-9 from top: Replace these lines with the following:

Multiplying these equations by $U_{r,m}$ for some integer $m \geq 1$ and then taking expectations it follows straightforwardly from (7.4), (7.5) and (7.10) that $\delta_m = 0$; hence, $W_t = \sum_{j=1}^{\infty} \beta_j W_{t-j}$ with probability 1. Q.E.D.

Page 185, line 3 of Th. 7.5: Replace $i$ with 0

Page 194, last line: Replace the leading + with −

Page 195, lines 4-6 from top: Replace these lines with the following statements:

Note that by Lemma 7.1,

$$|\exp(-x^2/2 + r(x))| = \exp(-x^2/2)|\exp(\text{Re}[r(x)] + i\text{Im}[r(x)])|$$

$$= \exp(-x^2/2)|\exp(\text{Re}[r(x)])| \leq \exp\left(-\frac{1}{2}x^2 + |x|^3\right) \leq 1 \text{ if } |x| \leq 1/2$$

so that

$$\left|\exp\left(-\frac{x^2}{2}(1/n)\sum_{t=1}^{n}X_t^2\right)\exp\left(\sum_{t=1}^{n}r\left(\frac{\xi X_t}{\sqrt{n}}\right)\right)\right| \leq \prod_{t=1}^{n} \left|\exp\left(-\frac{1}{2}\left(\frac{\xi X_t}{\sqrt{n}}\right)^2 + r\left(\frac{\xi X_t}{\sqrt{n}}\right)\right)\right|$$

$$\leq \sup_{|x| \leq |\xi|, \max_{1 \leq t \leq n}|X_t|/\sqrt{n}} |\exp(-x^2/2 + r(x))|^n \leq 1$$

if $|\xi|, \max_{1 \leq t \leq n}|X_t|/\sqrt{n} \leq 1/2$. Condition (7.38) implies that the latter event has probability converging to 1, hence pointwise in $\xi$,

$$\lim_{n \to \infty} P[|Z_n(\xi)| \leq 2] = 1.$$  (7.50)
Therefore, it follows from (7.49), (7.50) and the dominated convergence theorem that

Page 196, equation (7.58): Replace $\alpha^2$ with $\sigma^2$

Page 202, lines 14-16 from top: Delete the sentence starting with “In particular,...”

Page 203, line 1-3 from bottom: Replace these lines with the following statements:

In general, $\hat{X}_t$ is defined as the limit of the projection $\hat{X}_{t,n}$ of $X_t$ on the subspace spanned by $X_{t-1}, X_{t-2}, \ldots, X_{t-n}$, in the sense that $\lim_{n \to \infty} E[(\hat{X}_t - \hat{X}_{t,n})^2] = 0$, where $\hat{X}_{t,n}$ takes the form $\hat{X}_{t,n} = \sum_{j=1}^{n} \beta_{n,j} X_{t-j}$. Note that due to the covariance stationarity condition the coefficients $\beta_{n,j}$ do not depend on $t$. However, there is no guarantee that $\lim_{n \to \infty} \beta_{n,j}$ exists. Nevertheless, in the following it will be assumed that $\lim_{n \to \infty} \beta_{n,j} = \beta_j$ for each $j$, and that $\hat{X}_t$ takes the form $\hat{X}_t = \sum_{j=1}^{\infty} \beta_j X_{t-j}$.

Page 204, line 1 from top: Replace this line with:

Then the $\beta_j$’s are the solutions of the normal equations

Page 204, bottom: Add the following paragraph:

Remark: The general proof of the Wold decomposition is given in:
http://econ.la.psu.edu/~hbierens/WOLD.PDF.

Page 217: Replace the whole Assumption 8.2 with:

Assumption 8.2: $\lim_{n \to \infty} \sup_{\theta \in \Theta} n^{-1} \ln(\hat{L}_n(\theta)) - E[n^{-1} \ln(\hat{L}_n(\theta))]| = 0$ and $\lim_{n \to \infty} \sup_{\theta \in \Theta} |n^{-1} E[\ln(\hat{L}_n(\theta)/\hat{L}_n(\theta_0))] - \ell(\theta|\theta_0)| = 0$, where $\ell(\theta|\theta_0)$ is a continuous function of $\theta$ such that for arbitrarily small $\delta > 0$, $\sup_{\theta \in \Theta: \left|\theta - \theta_0\right| > \delta} \ell(\theta|\theta_0) < 0$.

Page 244, line 16 from top: Replace $P_{i,j}P_{j,i} = 1$ with $P_{i,j}P_{j,i} = I$
Page 253, line 1 of Section I.8: Replace “Theorem I.9” with “Theorem I.11”

Page 254, line 3 of Th. I.12: Replace “Theorem I.9” with “Theorem I.11”

Page 254, line 3 of Th. I.13: Replace “Theorem I.9” with “Theorem I.11”

Page 256, last line of Section I.8: Replace “Theorem I.5” with “Theorem I.15”

Page 265, lines 5-8 from bottom: This statement is incorrect. Replace these lines with:

Taking the transpose of a square matrix does not affect the determinant:

Page 265, lines 1-3 from bottom: Replace these lines with:

This result follows straightforwardly from Theorem I.20 and the following argument.

Page 266, line 3 from bottom: Replace \( \det(A) \) with \( \det(AB) \)

Page 268, line 2 from bottom: Replace "block-diagonal" with "block-triangular"

Page 269, line 2 from top: Replace \( k \times m \) with \( m \times k \)

Page 270, line 10 from top: Replace \( E_n \) with \( E_n \)

Page 270, line 11 from top: Insert the following text.

For example, in the 2×2 case

\[
A = \begin{pmatrix}
a_{1,1} & a_{1,2} \\
a_{2,1} & a_{2,2}
\end{pmatrix}
\]

we have

\[
\delta(A) = \det[p(A | 1, 2)] + \det[p(A | 2, 1)] = \det\left[\begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}\right] + \det\left[\begin{pmatrix} 0 & a_{1,2} \\ a_{2,1} & 0 \end{pmatrix}\right]
\]

\[
= a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = \det(A)
\]

Moreover, in the 3×3 case

\[
A = \begin{pmatrix}
a_{1,1} & a_{1,2} & a_{1,3} \\
a_{2,1} & a_{2,2} & a_{2,3} \\
a_{3,1} & a_{3,2} & a_{3,3}
\end{pmatrix}
\]
we have

$$\delta(A) = \det[\rho(A|1,2,3)] + \det[\rho(A|1,3,2)] + \det[\rho(A|2,1,3)]$$

$$+ \det[\rho(A|2,3,1)] + \det[\rho(A|3,1,2)] + \det[\rho(A|3,2,1)]$$

$$= \begin{vmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{vmatrix} + \begin{vmatrix} a_{1,1} & 0 & 0 \\ 0 & 0 & a_{2,3} \\ 0 & a_{3,2} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{1,2} & 0 \\ 0 & 0 & a_{2,3} \\ 0 & a_{3,2} & 0 \end{vmatrix}$$

$$+ \begin{vmatrix} 0 & a_{1,2} & 0 \\ 0 & 0 & a_{2,3} \\ a_{3,1} & 0 & 0 \end{vmatrix} + \begin{vmatrix} a_{2,1} & 0 & 0 \\ 0 & a_{2,3} & 0 \\ 0 & 0 & a_{3,3} \end{vmatrix} + \begin{vmatrix} a_{2,1} & 0 & 0 \\ a_{3,1} & 0 & 0 \\ a_{3,1} & a_{2,2} & 0 \end{vmatrix}$$

$$= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3} + a_{3,1}a_{2,1}a_{2,3} + a_{2,1}a_{3,2}a_{1,3} - a_{3,1}a_{2,2}a_{1,3}$$

$$= a_{1,1} \det \begin{pmatrix} a_{2,2} & a_{3,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} - a_{1,2} \det \begin{pmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{pmatrix} + a_{1,3} \det \begin{pmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix}.$$

The latter expression turns out to be the determinant of $A$, computed via the cofactor method. See Theorem I.31.

Page 270, line 7 from bottom: The equality $\rho(AB|i_1,i_2,\ldots,i_n) = A.\rho(B|i_1,i_2,\ldots,i_n)$ is incorrect, except if $A$ is a diagonal matrix. The following correction fixes this problem.

Page 270, lines 6-9 from bottom: Replace these lines with the following:

To check that $\delta(AB) = \delta(A)\delta(B)$, let in first instance $A$ be a lower or upper-triangular matrix with diagonal elements all equal to 1. Recall that these matrices can be written as products of elementary matrices. Therefore, if $\delta(AB) = \delta(B)$ for an arbitrary elementary matrix $A$ then $\delta(AB) = \delta(B)$ for triangular matrices with unit diagonal elements as well.

---

7 Thanks to Dr. Sang-bong Oh for bringing this problem to my attention.
For example, let
\[
A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix},
\]

hence
\[
AB = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} + c b_{3,1} & b_{2,2} + c b_{3,2} & b_{2,3} + c b_{3,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix}.
\]

Denote
\[
B_* = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix}.
\]

Then \( \det[p(AB|i_1,i_2,i_3)] = \det[p(B|i_1,i_2,i_3)] + c \det[p(B_*|i_1,i_2,i_3)] \), as is not hard to verify, hence \( \delta(AB) = \delta(B) + c \delta(B_*) \). However, swapping rows 2 and 3 of \( B_* \) it follows that \( \delta(B_*) = -\delta(B) \), hence \( \delta(B_*) = 0 \). Thus, in general, \( \delta(AB) = \delta(B) \) if \( A \) is a triangular matrix with unit diagonal elements.

Next, let \( A \) be a diagonal matrix, for example
\[
A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}.
\]

Then
\[
AB = \begin{pmatrix} a_1 b_{1,1} & a_1 b_{1,2} & a_1 b_{1,3} \\ a_2 b_{2,1} & a_2 b_{2,2} & a_2 b_{2,3} \\ a_3 b_{3,1} & a_3 b_{3,2} & a_3 b_{3,3} \end{pmatrix}.
\]
hence $\rho(AB|i_1,i_2,i_3) = A\rho(B|i_1,i_2,i_3)$ and thus $\delta(AB) = \det(A)\delta(B)$.

As has been shown before, the latter also holds if $A$ is a permutation matrix. It follows therefore from the decomposition $A = P^TLDU$ that the following result\(^8\) holds.

---

\(^8\) I.e., equation (1.53).

Page 286, line 19 from top: Replace $A \preceq B$ with $A \succeq C$