Appendix G
Basic Large Sample Theory

We define various quantities, and state results, that are useful in various places in the book. The presentation is informal see, for example, van der Vaart (1998) for more rigour.

Modes of Convergence

Suppose that Y_n, $n \geq 1$, are all random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ where Ω is a set (the sample space) $\mathcal{A}$ is a σ-algebra of subsets of Ω, and P is a probability measure.

Definition. We say that Y_n converges *almost surely* to Y, denoted $Y_n \rightarrow_{a.s.} Y$, if

$$Y_n(\omega) \rightarrow Y(\omega) \quad \text{for all } \omega \in A \text{ where } P(A^c) = 0 \tag{G.1}$$

or, equivalently, if, for every $\epsilon > 0$

$$P\left(\sup_{m \geq n} |Y_m - Y| > \epsilon\right) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \tag{G.2}$$

Definition. We say that Y_n converges *in probability* to Y, denoted $Y_n \rightarrow_p Y$, if

$$P\left(|Y_m - Y| > \epsilon\right) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \tag{G.3}$$

Definition. Define the distribution function of Y as $F(y) = \Pr(Y \leq y)$. We say that Y_n converges *in distribution* to Y, denoted $Y_n \rightarrow_d Y$, or $F_n \rightarrow F$, if

$$F_n(y) \rightarrow F(y) \quad \text{as} \quad n \rightarrow \infty \quad \text{for each continuity point } y \text{ of } F. \tag{G.4}$$

J. Wakefield, *Bayesian and Frequentist Regression Methods*, Springer Series in Statistics, DOI 10.1007/978-1-4419-0925-1_19,

Limit Theorems

Proposition (Weak Law of Large Numbers). *If $Y_1, Y_2, \ldots, Y_n, \ldots$ are independent and identically distributed (i.i.d.) with mean $\mu = \mathrm{E}[Y]$ (so $\mathrm{E}[|Y|] < \infty$) then $\overline{Y}_n \to_p \mu$.*

Proposition (Strong Law of Large Numbers). *If $Y_1, Y_2, \ldots, Y_n, \ldots$ are i.i.d. with mean $\mu = \mathrm{E}[Y]$ (so $\mathrm{E}[|Y|] < \infty$) then $\overline{Y}_n \to_{a.s.} \mu$.*

Proposition (Central Limit Theorem). *If $Y_1, Y_2, \ldots, Y_n$ are i.i.d. with mean $\mu = \mathrm{E}[Y]$ and variance σ^2 (so $\mathrm{E}[Y^2] < \infty$), then $\sqrt{n}(\overline{Y}_n - \mu) \to_d \mathrm{N}(0, \sigma^2)$.*

Proposition (Slutsky's Theorem). *Suppose that $A_n \to_p a$, $B_n \to_p b$, for constants a and b, and $Y_n \to_d Y$. Then $A_n Y_n + B_n \to_d aY + b$.*

Proposition (Delta Method). *Suppose $\sqrt{n}\ (\boldsymbol{Y}_n - \boldsymbol{\mu}) \to_d \boldsymbol{Z}$ and suppose that $\boldsymbol{g} : \mathbb{R}^p \to \mathbb{R}^k$ has a derivative $\boldsymbol{g}'$ at $\boldsymbol{\mu}$ (here $\boldsymbol{g}'$ is a $k \times p$ matrix of derivatives). Then the delta method gives the asymptotic distribution as*

$$\sqrt{n}\ [\boldsymbol{g}(\boldsymbol{Y}) - \boldsymbol{g}(\boldsymbol{\mu})] \to_d \boldsymbol{g}'(\boldsymbol{\mu})\boldsymbol{Z}.$$

If $\boldsymbol{Z} \sim \mathrm{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$, then

$$\sqrt{n}\ [\boldsymbol{g}(\boldsymbol{Y}) - \boldsymbol{g}(\boldsymbol{\mu})] \to_d \mathrm{N}_k \left[\mathbf{0}, \boldsymbol{g}'(\boldsymbol{\mu})\boldsymbol{\Sigma}\boldsymbol{g}'(\boldsymbol{\mu})^{\mathsf{T}}\right].$$