
CONTENTS

CHAPTER ONE PROBABILITY SPACES 1

Introduction 1

- 1.1. A probability space in which each simple event has probability zero. 3
- 1.2. A probability space in which each simple event has probability one. 4
- 1.3. Two events that are not disjoint but such that the probability of the union is equal to the sum of the probabilities. 4
- 1.4. An event that is independent of itself. 5
- 1.5. Two independent but disjoint events. 5
- 1.6. Three events that are pairwise independent but not independent. 6
- 1.7. N events that are pairwise independent but not independent. 6
- 1.8. An infinite sequence of events such that every pair is independent, but no triple is independent. 7
- 1.9. A sequence of any finite number of events, each with probability $\frac{1}{2}$, such that every proper subset is independent but the sequence itself is not. 8
- 1.10. For every $n \geq k \geq 2$ there are events B_1, \dots, B_n such that all the product rules hold except the one over $\{1, \dots, k\}$. 8
- 1.11. Three events such that the probability of the intersection is equal to the product of the probabilities even though the events are not independent. 9
- 1.12. N events such that the probability of the intersection is equal to the product of the probabilities even though the events are not independent. 10
- 1.13. Two independent events for which the conditional probability of A given B is not equal to the probability of A . 10
- 1.14. Simpson's Paradox: What's good for the population can be bad for every subgroup. 11
- 1.15. Events for which the sum of the probabilities is infinite, but only finitely many occur with probability one. 12
- 1.16. Independent events with probabilities less than one, yet whose infinite intersection occurs with nonzero probability. 12
- 1.17. A countable collection of sequences of events such that the probability tends to one for each sequence of events, yet the probability of the intersection over the collection does not tend to one. 13
- 1.18. A discrete probability measure on the real numbers whose support is the entire real line. 13

- 1.19. Two probability measures with the same support but that are not equivalent to each other. 14
- 1.20. A probability space in which no singleton sets $\{\omega\}$ are measurable. 14
- 1.21. The union of a collection of σ -fields that is not itself a σ -field. 15
- 1.22. A separable σ -field containing a nonseparable sub- σ -field. 15
- 1.23. A tail event with probability strictly between zero and one. 15
- 1.24. A tail σ -field formed from a nonindependent sequence of events, but for which every tail event has probability either zero or one. 16
- 1.25. Two probability measures defined on the same measurable space that agree on a class of sets but do not agree on the generated σ -field. 16
- 1.26. Two classes of events that are independent of each other, but whose generated σ -fields are dependent. 17
- 1.27. A σ -field \mathbf{G} generated by a class of sets \mathbf{A} [so that $\mathbf{G} = \sigma(\mathbf{A})$], but $\sigma(\{A_1 \times A_2 : A_1, A_2 \in \mathbf{A}\})$ is not equal to $\sigma(\{G_1 \times G_2 : G_1, G_2 \in \mathbf{G}\})$. 17
- 1.28. A finitely additive probability that is not countably additive; hence, a finitely additive probability on a field need not extend to a countably additive probability on the generated σ -field. 18
- 1.29. A nonmeasurable subset of the real numbers with respect to Lebesgue measure. 18
- 1.30. A probability space on the real numbers for which every subset is measurable. 20
- 1.31. A set of axioms under which every subset of the real numbers is Lebesgue-measurable. 20
- 1.32. An uncountable set with Lebesgue measure zero. 21

CHAPTER TWO RANDOM VARIABLES, DENSITIES, AND DISTRIBUTION FUNCTIONS 23

Introduction 23

- 2.1. A real-valued function, on a probability space, that is not a random variable. 26
- 2.2. A real random variable X and a real function Y both defined on a common probability space such that $P(X = Y) = 1$, but Y is not a real random variable. 27
- 2.3. A cumulative distribution function that is continuous but does not have a density. 27
- 2.4. A distribution that is neither discrete nor continuous. 28
- 2.5. An unbounded density function. 29
- 2.6. A discontinuous density function. 29
- 2.7. A continuous density that does not tend asymptotically to zero. 29
- 2.8. Two different densities that generate the same distribution. 30
- 2.9. An infinite family of bivariate distributions with the same marginal distributions. 30
- 2.10. A bivariate distribution that is not bivariate Gaussian but that has Gaussian marginal distributions. 31
- 2.11. Two Gaussian random variables whose sum does not have a Gaussian distribution. 32
- 2.12. Three pairwise independent, Gaussian random variables that are not trivariate Gaussian. 33
- 2.13. A bivariate distribution without a density but whose marginal distributions possess densities. 34
- 2.14. A discontinuous bivariate density with continuous marginal densities. 34
- 2.15. A continuous bivariate density with a discontinuous marginal density. 35

- 2.16.** A continuous function from \mathbf{R}^2 to $[0, 1]$, increasing in each coordinate separately and also satisfying $\lim_{x \rightarrow -\infty} F(x, y) = 0$ for all y , $\lim_{y \rightarrow -\infty} F(x, y) = 0$ for all x , and $\lim_{x, y \rightarrow \infty} F(x, y) = 1$, and yet F is not a distribution function on \mathbf{R}^2 . 35
- 2.17.** Two random variables that are real-valued functions of a real variable, but whose composition is not a random variable. 36

CHAPTER THREE MOMENTS OF RANDOM VARIABLES 38

Introduction 38

- 3.1.** A random variable with infinite mean. 39
- 3.2.** A random variable whose mean does not exist. 40
- 3.3.** Two random variables such that the expectation of the sum is not equal to the sum of the expectations. 41
- 3.4.** A finite expectation of a sum of three random variables that is not determined by the marginal distributions. 42
- 3.5.** Two random variables such that the expectation of their product is not the product of the expectations. 42
- 3.6.** A random variable such that the expectation of its reciprocal is not the reciprocal of its expectation. 43
- 3.7.** Two random variables such that the expectation of the minimum is not the minimum of the expectations. 44
- 3.8.** A random variable whose first moment exists but no higher moments exist. 44
- 3.9.** A random variable X and a positive number p such that $\lim_{x \rightarrow \infty} x^p P(|X| > x) = 0$, but $E(|X|^p) = \infty$. 45
- 3.10.** A random variable whose moment-generating function does not exist. 46
- 3.11.** A random variable, all of whose moments exist, but whose moment-generating function does not exist. 46
- 3.12.** A random variable whose first five moments are equal to those of a standard Gaussian, but which does not have a Gaussian distribution. 47
- 3.13.** A random variable whose first and third moments are both zero, but which is not symmetrically distributed. 47
- 3.14.** A distribution not determined uniquely by its moments. 48
- 3.15.** A class of distributions with identical moment sequences. 49
- 3.16.** A probability density that is not an even function even though all odd moments vanish. 49
- 3.17.** A distribution whose characteristic function is differentiable at zero but whose first moment does not exist. 50
- 3.18.** A random variable whose expectation does not exist, yet the expectations of certain order statistics from a sample do exist. 51

CHAPTER FOUR PROPERTIES OF REAL RANDOM VARIABLES 53

Introduction 53

- 4.1.** Two random variables that are equal in distribution, but whose difference is not equal in distribution to the zero distribution. 55
- 4.2.** Random variables such that the sum of the modes is not the mode of the sum. 55
- 4.3.** A random variable such that the reciprocal of the median is not a median of the reciprocal. 56

- 4.4. Random variables such that the median of the sum is not the sum of the medians. 57
- 4.5. Independent and identically distributed random variables such that each term has the same distribution as the sample mean. 58
- 4.6. Independent and identically distributed random variables such that the sample mean is more disperse than each term. 59
- 4.7. A convolution of two unimodal continuous distributions that is not unimodal. 60
- 4.8. A convolution of two singular distributions that is absolutely continuous. 61
- 4.9. Random variables that are pairwise independent but not independent. 62
- 4.10. An infinite sequence of random variables such that every pair is independent, but no triple is independent. 62
- 4.11. Two random variables that are uncorrelated but not independent. 63
- 4.12. Two random variables that are uncorrelated even though one may be predicted perfectly from the other. 63
- 4.13. Two independent random variables such that the expectation of the quotient is not the quotient of the expectations. 64
- 4.14. Two independent integrable random variables such that the quotient does not have an expectation even though it is always finite. 65
- 4.15. Two random variables for which the expectation of the product is equal to the product of the expectations, but they are not independent. 65
- 4.16. Uncorrelated Gaussian random variables that are not independent. 65
- 4.17. Random variables that are exchangeable but not independent. 66
- 4.18. Random variables that are independent but not exchangeable. 66
- 4.19. Random variables that are identically distributed but not exchangeable. 67
- 4.20. Random variables that are exchangeable but not uncorrelated. 67
- 4.21. Two marginal distributions for which no joint distribution exists with a correlation of -1 . 68
- 4.22. A family of bivariate distributions such that the range of possible correlations is a small subset of $[-1, 1]$. 68
- 4.23. A symmetric matrix with positive determinant that is not a covariance matrix. 69
- 4.24. A positive definite quadratic form of a multivariate Gaussian distribution that does not have a chi-squared distribution. 69
- 4.25. Random variables X , Y , and Z such that Y is positively correlated with both X and Z , but X and Z are negatively correlated with each other. 70
- 4.26. Random variables X , Y , and Z such that X is uncorrelated with both Y and Z , but Y is correlated with Z . 70
- 4.27. Random variables that are perfectly correlated, yet one is not a linear function of the other. 70
- 4.28. A function of two independent random variables, X and Y , that is independent of X but is not almost surely a function of Y alone. 71
- 4.29. Random variables such that the characteristic function of their sum is the product of their characteristic functions, yet they are not independent. 71
- 4.30. Random variables such that the density of their sum is the convolution of their densities, yet they are not independent. 72
- 4.31. Two distinct characteristic functions that are equal in an interval. 74
- 4.32. Two distinct characteristic functions whose absolute values are equal at all real values. 75
- 4.33. Convolutions $F * G$ and $F * H$ that are equal even though G and H are not equal. 76
- 4.34. Identically distributed random variables such that their difference does not have a symmetric distribution. 77

- 4.35. Two real random variables, each symmetric about zero, but whose sum is not symmetrically distributed. 77
- 4.36. Two random variables, each symmetric about zero, but whose sum is symmetric about a nonzero center. 78
- 4.37. A sum of independent and identically distributed random variables that is symmetric about zero, even though the individual distributions are not. 78
- 4.38. A characteristic function that is real in a neighborhood of zero but is not everywhere real. 80
- 4.39. A distribution that is symmetric but is not the result of a symmetrization procedure. 80
- 4.40. A characteristic function that never vanishes but is not infinitely divisible. 80
- 4.41. An infinitely divisible distribution that is not a stable law. 81
- 4.42. Random variables X and Y for which the best predictor of Y given X is not the inverse function of the best predictor of X given Y . 81
- 4.43. Random variables X and Y such that X predicts Y perfectly, but Y has no power in predicting X . 82
- 4.44. Nonconstant random variables X and Y such that Y is a function of X , but the best predictor of X given Y is zero. 82

CHAPTER FIVE SEQUENCES OF RANDOM VARIABLES 83

Introduction 83

- 5.1. A sequence of random variables that converges in law to each of two limiting random variables that are almost surely different. 88
- 5.2. Two independent sequences of random variables, each converging in law to a limit, such that the sequence of term-by-term sums does not converge in law to the sum of the limits. 88
- 5.3. Two sequences of random variables, each converging in law to a limit, such that the joint distribution does not converge in law to any limit. 89
- 5.4. A sequence of distribution functions that converges in distribution such that the sequence of expected values of a bounded function does not converge to the expectation of that same function under the limiting distribution. 90
- 5.5. A sequence of distribution functions that converges in distribution such that the sequence of expected values of a continuous function does not converge to the expectation of that same function under the limiting distribution. 90
- 5.6. A sequence of random variables that converges almost surely (and hence in probability) to a random variable X but does not converge in mean to X . 91
- 5.7. A sequence of random variables that converges in probability to a random variable X but does not converge almost surely to X . 91
- 5.8. A sequence of independent random variables that converges in probability to a random variable X but does not converge almost surely to X . 92
- 5.9. A sequence of independent random variables that converges in probability to a random variable, but the \limsup is almost surely ∞ and the \liminf is almost surely $-\infty$. 93
- 5.10. A sequence of random variables that converges almost surely to a limit, and a function such that the sequence of function values does not even converge in probability to the function evaluated at the limit. 94
- 5.11. Random variables X_n converging in law to X and a continuous measurable real-valued function g such that $g(X_n)$ does not converge in law to $g(X)$, even though the set of discontinuities of g has probability zero under the law of X . 94

- 5.12. A determining class that is not a convergence-determining class. 95
- 5.13. A sequence of random variables that converges in quadratic mean to a random variable X but does not converge almost surely to X . 95
- 5.14. A sequence of random variables that converges in the p th mean to a random variable X for every positive value of p but does not converge almost surely to X . 96
- 5.15. A sequence of random variables that converges in probability to a random variable X but does not converge in the p th mean. 96
- 5.16. A sequence of random variables that converges almost surely to X but does not converge in the p th mean to X . 97
- 5.17. A sequence of random variables that converges in distribution to X but not in probability to X . 97
- 5.18. A sequence of random variables that converges in distribution to X but does not converge in the p th mean to X . 98
- 5.19. A sequence of random variables whose characteristic functions converge, but the random variables do not converge in distribution to any random variable. 98
- 5.20. A sequence of random variables that converges in distribution but whose cumulative distribution functions do not converge pointwise. 100
- 5.21. A sequence of distributions F_n on the real line for which no subsequence converges weakly to a distribution, even though for every real number x , the sequence $F_n(x)$ converges. 100
- 5.22. A sequence of random variables that converges in distribution but whose sequence of corresponding probability measures does not converge for all Borel sets. 101
- 5.23. A sequence of discrete random variables that converges in distribution to a continuous random variable. 102
- 5.24. A sequence of continuous random variables that converges in distribution to a discrete random variable. 102
- 5.25. A sequence of absolutely continuous distributions with support equal to the entire plane that converges to a limit law that is degenerate at the origin. 103
- 5.26. A sequence of continuous random variables converging in distribution to a continuous random variable even though the sequence of densities does not converge. 104
- 5.27. A sequence of discrete random variables that converges in distribution to a discrete random variable even though the sequence of discrete probability functions does not converge. 104
- 5.28. A sequence of dependent bivariate random variables that converges in distribution to an independent bivariate random variable. 105
- 5.29. A sequence of discrete random variables that converges almost surely and in distribution but does not converge in the p th mean. 105
- 5.30. A sequence of continuous random variables with corresponding densities that converge pointwise but not in mean. 106
- 5.31. A sequence of nonsymmetric distributions that converges in distribution to a symmetric distribution. 106
- 5.32. An example where the sum of the expectations is not equal to the expectation of the sum. 106
- 5.33. A sequence of random variables that converges almost surely to a limit such that the expectations converge to the expectation of the limit, yet the variances do not converge to the variance of the limit. 107
- 5.34. A sequence of random variables with finite moments of all orders that converges in distribution, yet the variances do not converge to any finite number. 107
- 5.35. A sequence X_n of real random variables such that $\sum_{n=1}^{\infty} P(|X_n| > \varepsilon) = \infty$, but X_n converges almost surely to zero. 108

- 5.36. A sequence of independent and identically distributed continuous random variables where the laws of large numbers do not hold. 108
- 5.37. A sequence of independent and identically distributed discrete random variables where the laws of large numbers do not hold. 109
- 5.38. Independent and identically distributed random variables whose partial averages converge in probability to zero even though the expectation of each term does not exist. 109
- 5.39. Independent and identically distributed random variables such that the partial averages converge in probability to zero but do not converge almost surely. 111
- 5.40. A sequence of independent random variables that converges in probability to zero but whose partial averages do not converge in probability to zero. 112
- 5.41. The strong law of large numbers does not extend directly to triangular arrays. 112
- 5.42. A limit in distribution of partial sums of independent random variables that is not infinitely divisible. 114
- 5.43. A sequence of independent and identically distributed random variables where the Central Limit Theorem does not hold. 115
- 5.44. A sequence of independent and identically distributed random variables with finite mean and finite variance whose sequence of sample averages, standardized by subtracting the mean and dividing by the standard deviation of the sample average, is not asymptotically Gaussian. 115
- 5.45. A sequence of identically distributed, pairwise independent random variables with finite nonzero variance for which the Central Limit Theorem does not hold. 115
- 5.46. A triangular array that satisfies the Lindeberg condition but does not satisfy any Lyapounov condition. 116
- 5.47. A sequence of independent and identically distributed random variables with finite moments of all positive real orders less than two, but whose partial sums do not converge to a Gaussian distribution when normalized by subtracting the expectation and dividing by the square root of the sample size. 117
- 5.48. A sequence of independent and identically distributed random variables with infinite variance whose sequence of partial averages is nonetheless asymptotically Gaussian. 119
- 5.49. A triangular array, where in each row the variables are independent and identically distributed with mean zero and variance one, but whose normalized partial sums do not converge to a Gaussian distribution. 120
- 5.50. A triangular array, each row consisting of pairwise independent identically distributed Bernoulli random variables such that the expected number of successes in each row is one, but the limiting distribution is not the Poisson distribution with mean one. 121
- 5.51. A sequence of independent random variables whose partial sums, when standardized to have mean zero and variance one, converge to a Gaussian distribution but with variance less than one. 121
- 5.52. A precise formulation of the uniformly asymptotically negligible condition is important. 122
- 5.53. A uniformly asymptotically negligible triangular array of independent random variables whose normalized row sums converge weakly to a Gaussian distribution, but for which the Lindeberg condition fails to hold. 123
- 5.54. A sequence of random variables, a sequence of normalizing constants, and a continuous function, such that the normalized sequence is asymptotically standard Gaussian, but applying the same normalization to the function values does not result in convergence to a Gaussian distribution. 124

- 5.55. A triangular array of random variables with independent random variables in each row that is not uniformly asymptotically negligible, but whose asymptotic distribution of the row sums is nonetheless standard Gaussian. 124

CHAPTER SIX
CONDITIONAL EXPECTATION, MARTINGALES, AND
ALMOST SURE CONVERGENCE OF SUMS OF
INDEPENDENT RANDOM VARIABLES 126

Introduction 126

- 6.1. Random variables X , Y , and Z such that X is independent of Y and of Z , $E(X|Y) = E(X|Z) = \frac{1}{2}$ is degenerate and hence independent of Y and of Z , but $E[X|(Y, Z)]$ is not independent of (Y, Z) . 129
- 6.2. Random variables such that Z is independent of X , Z is independent of Y , but $E[X|(Y, Z)] \neq E(X|Y)$ almost surely. 129
- 6.3. $E(Y|X) = E(Y)$ need not imply that X and Y are independent, and $\text{Cov}(X, Y) = 0$ need not imply that $E(Y|X) = E(Y)$ almost surely. 130
- 6.4. Random variables X and Y such that $E(Y|X)$ exists, but Y is not integrable. 131
- 6.5. A Borel-measurable function f and random variables U, V such that $E[f(U, V)|U]$ is almost surely equal to one, but for any fixed u we have $E[f(u, V)|U]$ is almost surely equal to zero. 132
- 6.6. Random variables X and Y defined on a probability space (Ω, \mathbf{F}, P) with \mathbf{G} a sub- σ -field of \mathbf{F} and $X = E(Y|\mathbf{G})$ almost surely, but X is not a version of the conditional expectation of Y given \mathbf{G} . 133
- 6.7. Independent random variables X and Y defined on (Ω, \mathbf{F}, P) and a sub- σ -field \mathbf{G} of \mathbf{F} , such that $E(X|\mathbf{G})$ and $E(Y|\mathbf{G})$ are not independent for any versions of these conditional expectations. 133
- 6.8. Random variables U, V , and W such that there exist versions of $E(W|U)$ and $E(W|V)$ that are continuous functions of U and V , respectively, the events $\{U = 1\}$ and $\{V = 0\}$ are the same, yet the conditional expectations differ when $U = 1$ and $V = 0$. 133
- 6.9. A random variable X and a sub- σ -field \mathbf{G} containing all one-point sets, yet $E(X|\mathbf{G})$ is not almost surely equal to X . 134
- 6.10. For every z , $T(X, z)$ has distribution Q and is independent of Y , but $T(X, Z(Y))$ does not have distribution Q and is not independent of Y . 135
- 6.11. A sequence X_1, X_2, \dots of random variables and a sub- σ -field \mathbf{B} such that $E(X_n|\mathbf{B}) \rightarrow E(X|\mathbf{B})$ in p th mean for all $p \geq 1$, but X_n does not converge in p th mean to X for any $p \geq 1$. Also, a sequence X_1, X_2, \dots of random variables such that X_n does converge in p th mean to X for all $p > 0$, but $E(X_n|\mathbf{B})$ does not converge to $E(X|\mathbf{B})$ almost surely. 136
- 6.12. A sequence X_1, X_2, \dots of uniformly integrable random variables and a σ -field \mathbf{B} such that X_n converges almost surely to X , but $E(X_n|\mathbf{B})$ does not converge almost surely to $E(X|\mathbf{B})$. 136
- 6.13. A situation in which a regular conditional probability distribution does not exist. 138
- 6.14. Martingales, defined on a common probability space, whose sum is not a martingale. 139
- 6.15. A convex function of a submartingale that is not a submartingale. 140
- 6.16. A convex function of a martingale that is not a submartingale. 141
- 6.17. A martingale X_1, X_2, \dots and an integrable stopping time T such that $E(X_T)$, the expected value of the stopped martingale at time T , is infinite. 141

- 6.18. Independent random variables D_1, D_2, \dots with mean zero and an integrable stopping time T for which the expectation of the sum of the first T terms is not zero; that is, $E(D_1 + \dots + D_T) \neq E(D_1) \cdot E(T)$. 142
- 6.19. A martingale X_1, X_2, \dots and an integrable stopping time T such that $E(|X_T|)$ is finite, but $E(X_1) \neq E(X_T)$. 143
- 6.20. A martingale X_1, X_2, \dots and an integrable stopping time T that satisfy $\liminf E[|X_n|I(T > n)] = 0$, but $E(X_1) \neq E(X_T)$. 143
- 6.21. A martingale, for which each term is a product of positive independent random variables with mean one, that converges almost surely, but for which the expectations do not converge to the expectation of the limit. 144
- 6.22. A martingale that tends to $-\infty$ with probability one. 145
- 6.23. A martingale that converges in probability but not almost surely. 146
- 6.24. A martingale that converges almost surely even though the supremum of the expected absolute values is infinite. 148
- 6.25. Nonuniqueness of last elements for sub- or supermartingales. 149
- 6.26. A supermartingale with a last element that is not uniformly integrable. 149
- 6.27. A sequence of independent random variables with zero means and finite variances whose partial sums converge almost surely even though the sum of the variances is infinite. 150
- 6.28. Independent random variables Y_1, Y_2, \dots such that $\sum P\{|Y_n| > 1\} < \infty$ and $\sum \text{Var}\{Y_n I(|Y_n| \leq 1)\} < \infty$, but for which $\sum Y_n$ diverges almost surely. 150
- 6.29. Independent random variables Y_1, Y_2, \dots such that $\sum E\{Y_n I(|Y_n| \leq 1)\}$ converges and $\sum \text{Var}\{Y_n I(|Y_n| < 1)\} < \infty$, but for which $\sum Y_n$ diverges almost surely. 151
- 6.30. Independent random variables Y_1, Y_2, \dots such that $\sum P(|Y_n| > 1) < \infty$ and $\sum E\{Y_n I(|Y_n| \leq 1)\}$ converges, but for which $\sum Y_n$ diverges almost surely. 151
- 6.31. A sequence Y_1, Y_2, \dots of independent random variables, each with finite mean and variance, and positive constants b_1, b_2, \dots such that $\sum [\text{Var}(Y_n)]/b_n^2 < \infty$ but $\sum_{j=1}^n [Y_j - E(Y_j)]/b_n$ does not converge almost surely to zero. 152
- 6.32. A sequence Y_1, Y_2, \dots of independent random variables, each with finite mean and variance, and nondecreasing positive constants b_1, b_2, \dots such that $b_n \rightarrow \infty$, satisfying $\sum_{j=1}^n [Y_j - E(Y_j)]/b_n \rightarrow 0$ almost surely, but $\sum \text{Var}(Y_n)/b_n^2 = \infty$. 153

CHAPTER SEVEN

PROPERTIES OF STATISTICAL EXPERIMENTS 154

Introduction 154

- 7.1. A one-parameter family of distributions for which no single continuous sufficient statistic exists. 158
- 7.2. A real-valued sufficient statistic for a sample from a Gaussian distribution with unknown mean and unknown variance. 159
- 7.3. A one-parameter family of distributions for which the order statistics are minimal sufficient. 160
- 7.4. A sufficient statistic (X, C) for a parameter θ such that C does not depend on θ ; that is, C is ancillary, yet X alone is not sufficient for θ . 161
- 7.5. Nonuniqueness of ancillary statistics. 162
- 7.6. Nonuniqueness of sufficient statistics. 163
- 7.7. A statistic (S, T) that is sufficient for (θ, ϕ) such that S is sufficient for θ when ϕ is fixed and known, but T is not sufficient for ϕ when θ is fixed and known. 163
- 7.8. Two statistics that are each ancillary but jointly are not ancillary. 164

- 7.9. An ancillary statistic that is not location invariant in a location family. 164
- 7.10. Unbiased estimators need not be unique. 165
- 7.11. A unique unbiased estimator. 166
- 7.12. A parameter for which no unbiased estimator exists. 167
- 7.13. A statistic that is unbiased for a parameter but is not median unbiased for that parameter. 167
- 7.14. A statistic that is median unbiased for a parameter but is not unbiased for that parameter. 168
- 7.15. A “ridiculous” unique unbiased estimator. 168
- 7.16. A statistic that is unbiased for a parameter, but whose square is not unbiased for the square of the parameter. 169
- 7.17. Unbiasedness is not an invariant property. 170
- 7.18. A statistic that is not complete. 170
- 7.19. A family of distributions that is not complete. 170
- 7.20. A discrete minimal sufficient statistic that is not complete. 171
- 7.21. A continuous minimal sufficient statistic that is not complete. 171
- 7.22. A statistic that is boundedly complete but not complete. 172
- 7.23. A minimal sufficient statistic that is not boundedly complete. 173
- 7.24. A family of univariate distributions indexed by a parameter (θ, ϕ) such that the family is complete with respect to θ when ϕ is known but is not even boundedly complete with respect to ϕ when θ is known. 173
- 7.25. In an exponential family, a natural minimal sufficient statistic that is not complete. 174

CHAPTER EIGHT CONSTRUCTION OF ESTIMATORS 175

Introduction 175

- 8.1. A biased method of moments estimator. 176
- 8.2. Nonuniqueness of method of moments estimators. 177
- 8.3. A method of moments estimator that may produce a value that could not be the true parameter. 178
- 8.4. A method of moments estimator that is not a function of complete sufficient statistics. 178
- 8.5. Nonexistence of a method of moments estimator. 179
- 8.6. An uninformative likelihood. 179
- 8.7. A biased maximum likelihood estimator. 180
- 8.8. Two data sets that each provide the same value of a maximum likelihood estimate but, when combined into one data set, provide a different value. 180
- 8.9. Nonuniqueness of solutions to the likelihood equation. 181
- 8.10. Nonuniqueness of the maximum likelihood estimator. 182
- 8.11. Nonexistence of a maximum likelihood estimator. 182
- 8.12. A maximum likelihood estimator that is discontinuous. 184
- 8.13. A maximum likelihood estimator that is not a function of a given sufficient statistic. 185
- 8.14. A maximum likelihood estimator that does not satisfy the likelihood equation. 186
- 8.15. A likelihood equation that yields a spurious solution for a maximum likelihood estimator. 186
- 8.16. A likelihood equation with a unique root, but no maximum likelihood estimator exists. 187

- 8.17. A maximum likelihood estimator that either does not satisfy the likelihood equation or is not unique. 188
- 8.18. A maximum likelihood estimator of the population mean that is not the sample mean. 189

CHAPTER NINE OPTIMALITY OF ESTIMATORS 190

Introduction 190

- 9.1. A uniformly minimum variance unbiased estimator whose variance does not attain the Cramer-Rao lower bound. 194
- 9.2. A “poor” uniformly minimum variance unbiased estimator. 195
- 9.3. A parameter for which no uniformly minimum variance unbiased estimator exists. 196
- 9.4. T is the uniformly minimum variance unbiased estimator for θ , but $q(T)$ is not the uniformly minimum variance unbiased estimator for $q(\theta)$. 196
- 9.5. A uniformly minimum variance unbiased estimator exists for θ but not for a function of θ . 197
- 9.6. Nonexistence of a uniformly minimum variance unbiased estimator for a discrete distribution even though an infinite class of unbiased estimators exists. 197
- 9.7. Nonexistence of a uniformly minimum variance unbiased estimator for a continuous distribution even though an infinite class of unbiased estimators exists. 199
- 9.8. An inefficient uniformly minimum variance unbiased estimator. 200
- 9.9. A Pitman estimator that has smaller mean squared error loss than the uniformly minimum variance unbiased estimator. 201
- 9.10. A Pitman estimator of location that is neither admissible nor minimax. 203
- 9.11. A minimum risk equivariant estimator that is neither minimax nor admissible. 203
- 9.12. A problem that remains invariant under two different groups, each leading to different minimum risk equivariant estimators. 204
- 9.13. A minimum risk equivariant estimator that is not risk unbiased. 204
- 9.14. The completeness condition cannot be omitted in the Lehmann-Scheffe Theorem. 205
- 9.15. Bounded completeness cannot replace completeness as a condition in the Lehmann-Scheffe Theorem. 205
- 9.16. Regularity conditions cannot be omitted in the Information Inequality. 206
- 9.17. A randomized minimax estimator. 207
- 9.18. An inadmissible method of moments estimator. 207
- 9.19. An inadmissible maximum likelihood estimator and the James-Stein Rule. 208
- 9.20. An observation from a one-parameter exponential family such that the natural sufficient statistic is inadmissible for its expectation. 209
- 9.21. A situation in which the mean squared error of the maximum likelihood estimator is always worse than using a constant estimator. 210
- 9.22. A situation where the maximum likelihood estimator and the uniformly minimum variance unbiased estimator differ and both are inadmissible. 211
- 9.23. A uniformly minimum variance unbiased estimator that renders the maximum likelihood estimator inadmissible. 213
- 9.24. An admissible uniformly minimum variance unbiased estimator that is not minimax. 213
- 9.25. An admissible estimator that is a pointwise and L^2 limit of inadmissible estimators. 215

- 9.26. Nonexistence of a Bayes estimator. 215
- 9.27. A pointwise and L^2 limit of unique Bayes estimators that is not Bayes. 216
- 9.28. An admissible estimator that is not Bayes. 216
- 9.29. A Bayes estimator that has constant risk, so that it is minimax, but it is inadmissible. 217
- 9.30. An inadmissible estimator that is a pointwise and L^2 limit of unique Bayes admissible estimators. 217
- 9.31. A minimax uniformly minimum variance unbiased estimator that is inadmissible. 218
- 9.32. A constant risk minimax estimator that is inadmissible. 219
- 9.33. A statistic that is asymptotically unbiased but not unbiased for any finite sample size. 219
- 9.34. A consistent estimator that is not unbiased for any finite sample size. 220
- 9.35. A consistent estimator that is asymptotically biased. 220
- 9.36. A unique maximum likelihood estimator that is not asymptotically Gaussian. 221
- 9.37. A unique maximum likelihood estimator whose rate of convergence is not the square root of n . 221
- 9.38. An example where S^2 and $[(n - 1)/n]S^2$ have different asymptotic distributions. 222
- 9.39. A maximum likelihood estimator that is asymptotically biased. 222
- 9.40. A Fisher efficient sequence of estimators whose bias and variance tend to infinity. 222
- 9.41. An inconsistent sequence of estimators that are unbiased for any finite sample size. 224
- 9.42. Nonuniqueness of consistent estimators. 224
- 9.43. A useless consistent estimator. 225
- 9.44. An inconsistent maximum likelihood estimator. 225
- 9.45. A family of distinct distributions with density functions $f(x; \theta)$ that are continuous in θ for all x , yet the maximum likelihood estimator for a sample of size n converges almost surely to one, regardless of the true value of θ . 227
- 9.46. An inconsistent maximum likelihood estimator for a function of the location parameter of a Gaussian distribution. 229
- 9.47. A superefficient estimator. 229
- 9.48. An inefficient method of moments estimator. 230

CHAPTER TEN HYPOTHESIS TESTING 232

Introduction 232

- 10.1. An unreasonable level 0.05 test result. 238
- 10.2. Two test statistics that differ almost surely, but nonetheless generate tests with identical power functions. 238
- 10.3. A hypothesis-testing problem where no nonrandomized tests at level less than 0.5 have any power. 239
- 10.4. A hypothesis-testing problem where the nonrandomized likelihood procedures have achievable levels of significance of only one and zero. 239
- 10.5. Nonexistence of a uniformly most powerful test for a simple hypothesis against only two alternatives. 240
- 10.6. Nonexistence of a uniformly most powerful test for a simple hypothesis against a one-sided alternative. 241

- 10.7. A most powerful one-sided test for a location parameter that does not reject for large values. 243
- 10.8. A family of distributions that is stochastically increasing but does not have monotone likelihood ratio. 245
- 10.9. A family of distributions with monotone likelihood ratio that is not an exponential family. 246
- 10.10. A location family that does not have monotone likelihood ratio, but a uniformly most powerful one-sided test exists. 246
- 10.11. An even density that generates a scale family with monotone likelihood ratio, but whose location family does not have monotone likelihood ratio. 247
- 10.12. An example where an optimal nonrandomized solution in testing a simple hypothesis against a simple alternative does not reject at the largest likelihood ratio. 247
- 10.13. Nonuniqueness of uniformly most powerful tests. 249
- 10.14. A uniformly most powerful test that does *not* depend on the data only through the complete sufficient statistic. 249
- 10.15. A uniformly most powerful two-sided test of a real parameter. 250
- 10.16. Most powerful tests of varying sizes with critical regions that are not nested. 251
- 10.17. Nonuniqueness of a least favorable distribution. 253
- 10.18. The least favorable distribution may depend on the size of the test. 253
- 10.19. Nonexistence of a uniformly most powerful unbiased test in a two-parameter exponential family. 254
- 10.20. A hypothesis-testing problem where a uniformly most powerful test exists conditional on the value of an ancillary statistic, but not even a uniformly most powerful unbiased unconditional test exists. 255
- 10.21. Existence of a uniformly most powerful unbiased test when a nuisance parameter is unknown even though no such optimal property holds when the nuisance parameter is known. 256
- 10.22. An inadmissible likelihood ratio test. 257
- 10.23. Nonuniqueness of a Bayes classification rule. 258
- 10.24. Nonuniqueness of a maximal invariant. 259
- 10.25. A test whose power function is invariant, but the test is not. 259
- 10.26. A uniformly most powerful invariant test that is not a likelihood ratio test. 260
- 10.27. An inadmissible uniformly most powerful invariant test. 260
- 10.28. A problem invariant under two groups such that the uniformly most powerful invariant tests are different under each group. 261
- 10.29. An almost invariant test that is not equivalent to an invariant test. 262
- 10.30. A uniformly most powerful invariant test that is not uniformly most powerful almost invariant. 263
- 10.31. A hypothesis-testing problem where both uniformly most powerful unbiased and uniformly most powerful invariant tests exist, but they do not coincide. 264
- 10.32. Nonexistence of a uniformly most powerful invariant test, even though a uniformly most powerful unbiased test exists. 264
- 10.33. A uniformly most powerful invariant test that is not maximin. 266

APPENDIX ADDITIONAL COUNTEREXAMPLES 267

- A.1. Probability and measure. 267
- A.2. Random variables and distributions. 268
- A.3. Independence. 269

A.4. Sequences of random variables. 269
A.5. Stochastic processes and martingales. 270
A.6. Estimation. 271
A.7. Sufficiency. 272
A.8. Likelihood procedures. 273
A.9. Bayes methods. 273
A.10. Admissibility. 274
A.11. Hypothesis testing. 275
A.12. Confidence intervals. 276
A.13. Categorical data analysis. 276
A.14. Miscellaneous. 277

References 278

Index 287