

# Stat 5101 Lecture Notes

Charles J. Geyer

Copyright 1998, 1999, 2000, 2001 by Charles J. Geyer

May 7, 2001



# Contents

<b>1</b>	<b>Random Variables and Change of Variables</b>	<b>1</b>
1.1	Random Variables . . . . .	1
1.1.1	Variables . . . . .	1
1.1.2	Functions . . . . .	1
1.1.3	Random Variables: Informal Intuition . . . . .	3
1.1.4	Random Variables: Formal Definition . . . . .	3
1.1.5	Functions of Random Variables . . . . .	7
1.2	Change of Variables . . . . .	7
1.2.1	General Definition . . . . .	7
1.2.2	Discrete Random Variables . . . . .	9
1.2.3	Continuous Random Variables . . . . .	12
1.3	Random Vectors . . . . .	14
1.3.1	Discrete Random Vectors . . . . .	15
1.3.2	Continuous Random Vectors . . . . .	15
1.4	The Support of a Random Variable . . . . .	17
1.5	Joint and Marginal Distributions . . . . .	18
1.6	Multivariable Change of Variables . . . . .	22
1.6.1	The General and Discrete Cases . . . . .	22
1.6.2	Continuous Random Vectors . . . . .	22
<b>2</b>	<b>Expectation</b>	<b>31</b>
2.1	Introduction . . . . .	31
2.2	The Law of Large Numbers . . . . .	32
2.3	Basic Properties . . . . .	32
2.3.1	Axioms for Expectation (Part I) . . . . .	32
2.3.2	Derived Basic Properties . . . . .	34
2.3.3	Important Non-Properties . . . . .	36
2.4	Moments . . . . .	37
2.4.1	First Moments and Means . . . . .	38
2.4.2	Second Moments and Variances . . . . .	40
2.4.3	Standard Deviations and Standardization . . . . .	42
2.4.4	Mixed Moments and Covariances . . . . .	43
2.4.5	Exchangeable Random Variables . . . . .	50
2.4.6	Correlation . . . . .	50

2.5	Probability Theory as Linear Algebra . . . . .	55
2.5.1	The Vector Space $L^1$ . . . . .	56
2.5.2	Two Notions of Linear Functions . . . . .	58
2.5.3	Expectation on Finite Sample Spaces . . . . .	59
2.5.4	Axioms for Expectation (Part II) . . . . .	62
2.5.5	General Discrete Probability Models . . . . .	64
2.5.6	Continuous Probability Models . . . . .	66
2.5.7	The Trick of Recognizing a Probability Density . . . . .	68
2.5.8	Probability Zero . . . . .	68
2.5.9	How to Tell When Expectations Exist . . . . .	70
2.5.10	$L^p$ Spaces . . . . .	74
2.6	Probability is a Special Case of Expectation . . . . .	75
2.7	Independence . . . . .	77
2.7.1	Two Definitions . . . . .	77
2.7.2	The Factorization Criterion . . . . .	77
2.7.3	Independence and Correlation . . . . .	78
<b>3</b>	<b>Conditional Probability and Expectation</b>	<b>83</b>
3.1	Parametric Families of Distributions . . . . .	83
3.2	Conditional Probability Distributions . . . . .	86
3.3	Axioms for Conditional Expectation . . . . .	88
3.3.1	Functions of Conditioning Variables . . . . .	88
3.3.2	The Regression Function . . . . .	89
3.3.3	Iterated Expectations . . . . .	91
3.4	Joint, Conditional, and Marginal . . . . .	95
3.4.1	Joint Equals Conditional Times Marginal . . . . .	95
3.4.2	Normalization . . . . .	97
3.4.3	Renormalization . . . . .	98
3.4.4	Renormalization, Part II . . . . .	101
3.4.5	Bayes Rule . . . . .	103
3.5	Conditional Expectation and Prediction . . . . .	105
<b>4</b>	<b>Parametric Families of Distributions</b>	<b>111</b>
4.1	Location-Scale Families . . . . .	111
4.2	The Gamma Distribution . . . . .	115
4.3	The Beta Distribution . . . . .	117
4.4	The Poisson Process . . . . .	119
4.4.1	Spatial Point Processes . . . . .	119
4.4.2	The Poisson Process . . . . .	120
4.4.3	One-Dimensional Poisson Processes . . . . .	122
<b>5</b>	<b>Multivariate Theory</b>	<b>127</b>
5.1	Random Vectors . . . . .	127
5.1.1	Vectors, Scalars, and Matrices . . . . .	127
5.1.2	Random Vectors . . . . .	128
5.1.3	Random Matrices . . . . .	128

5.1.4	Variance Matrices . . . . .	129
5.1.5	What is the Variance of a Random Matrix? . . . . .	130
5.1.6	Covariance Matrices . . . . .	131
5.1.7	Linear Transformations . . . . .	133
5.1.8	Characterization of Variance Matrices . . . . .	135
5.1.9	Degenerate Random Vectors . . . . .	136
5.1.10	Correlation Matrices . . . . .	140
5.2	The Multivariate Normal Distribution . . . . .	141
5.2.1	The Density . . . . .	143
5.2.2	Marginals . . . . .	146
5.2.3	Partitioned Matrices . . . . .	146
5.2.4	Conditionals and Independence . . . . .	148
5.3	Bernoulli Random Vectors . . . . .	151
5.3.1	Categorical Random Variables . . . . .	152
5.3.2	Moments . . . . .	153
5.4	The Multinomial Distribution . . . . .	154
5.4.1	Categorical Random Variables . . . . .	154
5.4.2	Moments . . . . .	155
5.4.3	Degeneracy . . . . .	155
5.4.4	Density . . . . .	156
5.4.5	Marginals and “Sort Of” Marginals . . . . .	157
5.4.6	Conditionals . . . . .	159
<b>6</b>	<b>Convergence Concepts</b>	<b>165</b>
6.1	Univariate Theory . . . . .	165
6.1.1	Convergence in Distribution . . . . .	165
6.1.2	The Central Limit Theorem . . . . .	166
6.1.3	Convergence in Probability . . . . .	169
6.1.4	The Law of Large Numbers . . . . .	170
6.1.5	The Continuous Mapping Theorem . . . . .	170
6.1.6	Slutsky’s Theorem . . . . .	171
6.1.7	Comparison of the LLN and the CLT . . . . .	172
6.1.8	Applying the CLT to Addition Rules . . . . .	172
6.1.9	The Cauchy Distribution . . . . .	174
<b>7</b>	<b>Sampling Theory</b>	<b>177</b>
7.1	Empirical Distributions . . . . .	177
7.1.1	The Mean of the Empirical Distribution . . . . .	179
7.1.2	The Variance of the Empirical Distribution . . . . .	179
7.1.3	Characterization of the Mean . . . . .	180
7.1.4	Review of Quantiles . . . . .	180
7.1.5	Quantiles of the Empirical Distribution . . . . .	181
7.1.6	The Empirical Median . . . . .	183
7.1.7	Characterization of the Median . . . . .	183
7.2	Samples and Populations . . . . .	185
7.2.1	Finite Population Sampling . . . . .	185

7.2.2	Repeated Experiments . . . . .	188
7.3	Sampling Distributions of Sample Moments . . . . .	188
7.3.1	Sample Moments . . . . .	188
7.3.2	Sampling Distributions . . . . .	190
7.3.3	Moments . . . . .	194
7.3.4	Asymptotic Distributions . . . . .	196
7.3.5	The $t$ Distribution . . . . .	199
7.3.6	The $F$ Distribution . . . . .	202
7.3.7	Sampling Distributions Related to the Normal . . . . .	202
7.4	Sampling Distributions of Sample Quantiles . . . . .	205
<b>8</b>	<b>Convergence Concepts Continued</b>	<b>211</b>
8.1	Multivariate Convergence Concepts . . . . .	211
8.1.1	Convergence in Probability to a Constant . . . . .	211
8.1.2	The Law of Large Numbers . . . . .	212
8.1.3	Convergence in Distribution . . . . .	212
8.1.4	The Central Limit Theorem . . . . .	214
8.1.5	Slutsky and Related Theorems . . . . .	215
8.2	The Delta Method . . . . .	218
8.2.1	The Univariate Delta Method . . . . .	218
8.2.2	The Multivariate Delta Method . . . . .	221
8.2.3	Asymptotics for Sample Moments . . . . .	223
8.2.4	Asymptotics of Independent Sequences . . . . .	224
8.2.5	Asymptotics of Sample Quantiles . . . . .	225
<b>9</b>	<b>Frequentist Statistical Inference</b>	<b>231</b>
9.1	Introduction . . . . .	231
9.1.1	Inference . . . . .	231
9.1.2	The Sample and the Population . . . . .	231
9.1.3	Frequentist versus Bayesian Inference . . . . .	232
9.2	Models, Parameters, and Statistics . . . . .	233
9.2.1	Parametric Models . . . . .	233
9.2.2	Nonparametric Models . . . . .	236
9.2.3	Semiparametric Models . . . . .	237
9.2.4	Interest and Nuisance Parameters . . . . .	237
9.2.5	Statistics . . . . .	237
9.3	Point Estimation . . . . .	238
9.3.1	Bias . . . . .	239
9.3.2	Mean Squared Error . . . . .	242
9.3.3	Consistency . . . . .	242
9.3.4	Asymptotic Normality . . . . .	243
9.3.5	Method of Moments Estimators . . . . .	244
9.3.6	Relative Efficiency . . . . .	249
9.3.7	Asymptotic Relative Efficiency (ARE) . . . . .	250
9.4	Interval Estimation . . . . .	251
9.4.1	Exact Confidence Intervals for Means . . . . .	252

9.4.2	Pivotal Quantities . . . . .	253
9.4.3	Approximate Confidence Intervals for Means . . . . .	254
9.4.4	Paired Comparisons . . . . .	256
9.4.5	Independent Samples . . . . .	256
9.4.6	Confidence Intervals for Variances . . . . .	264
9.4.7	The Role of Asymptotics . . . . .	266
9.4.8	Robustness . . . . .	269
9.5	Tests of Significance . . . . .	272
9.5.1	Interest and Nuisance Parameters Revisited . . . . .	275
9.5.2	Statistical Hypotheses . . . . .	275
9.5.3	Tests of Equality-Constrained Null Hypotheses . . . . .	276
9.5.4	$P$ -values . . . . .	280
9.5.5	One-Tailed Tests . . . . .	282
9.5.6	The Duality of Tests and Confidence Intervals . . . . .	285
9.5.7	Sample Size Calculations . . . . .	287
9.5.8	Multiple Tests and Confidence Intervals . . . . .	292
<b>10</b>	<b>Likelihood Inference</b>	<b>303</b>
10.1	Likelihood . . . . .	303
10.2	Maximum Likelihood . . . . .	305
10.3	Sampling Theory . . . . .	308
10.3.1	Derivatives of the Log Likelihood . . . . .	308
10.3.2	The Sampling Distribution of the MLE . . . . .	314
10.3.3	Asymptotic Relative Efficiency . . . . .	317
10.3.4	Estimating the Variance . . . . .	319
10.3.5	Tests and Confidence Intervals . . . . .	320
10.4	Multiparameter Models . . . . .	325
10.4.1	Maximum Likelihood . . . . .	326
10.4.2	Sampling Theory . . . . .	330
10.4.3	Likelihood Ratio Tests . . . . .	336
10.5	Change of Parameters . . . . .	342
10.5.1	Invariance of Likelihood . . . . .	342
10.5.2	Invariance of the MLE . . . . .	343
10.5.3	Invariance of Likelihood Ratio Tests . . . . .	344
10.5.4	Covariance of Fisher Information . . . . .	344
<b>11</b>	<b>Bayesian Inference</b>	<b>349</b>
11.1	Parametric Models and Conditional Probability . . . . .	349
11.2	Prior and Posterior Distributions . . . . .	350
11.2.1	Prior Distributions . . . . .	350
11.2.2	Posterior Distributions . . . . .	352
11.3	The Subjective Bayes Philosophy . . . . .	357
11.4	More on Prior and Posterior Distributions . . . . .	359
11.4.1	Improper Priors . . . . .	359
11.4.2	Conjugate Priors . . . . .	361
11.4.3	The Two-Parameter Normal Distribution . . . . .	362

11.5	Bayesian Point Estimates . . . . .	367
11.6	Highest Posterior Density Regions . . . . .	369
11.7	Bayes Tests . . . . .	372
11.8	Bayesian Asymptotics . . . . .	376
<b>12</b>	<b>Regression</b>	<b>381</b>
12.1	The Population Regression Function . . . . .	381
12.1.1	Regression and Conditional Expectation . . . . .	381
12.1.2	Best Prediction . . . . .	382
12.1.3	Best Linear Prediction . . . . .	383
12.2	The Sample Regression Function . . . . .	387
12.3	Sampling Theory . . . . .	391
12.3.1	The Regression Model . . . . .	391
12.3.2	The Gauss-Markov Theorem . . . . .	393
12.3.3	The Sampling Distribution of the Estimates . . . . .	396
12.3.4	Tests and Confidence Intervals for Regression Coefficients	398
12.3.5	The Hat Matrix . . . . .	400
12.3.6	Polynomial Regression . . . . .	402
12.3.7	The $F$ -Test for Model Comparison . . . . .	406
12.3.8	Intervals for the Regression Function . . . . .	410
12.4	The Abstract View of Regression . . . . .	413
12.5	Categorical Predictors (ANOVA) . . . . .	416
12.5.1	Categorical Predictors and Dummy Variables . . . . .	416
12.5.2	ANOVA . . . . .	423
12.6	Residual Analysis . . . . .	424
12.6.1	Leave One Out . . . . .	426
12.6.2	Quantile-Quantile Plots . . . . .	430
12.7	Model Selection . . . . .	433
12.7.1	Overfitting . . . . .	434
12.7.2	Mean Square Error . . . . .	436
12.7.3	The Bias-Variance Trade-Off . . . . .	438
12.7.4	Model Selection Criteria . . . . .	438
12.7.5	All Subsets Regression . . . . .	442
12.8	Bernoulli Regression . . . . .	450
12.8.1	A Dumb Idea (Identity Link) . . . . .	450
12.8.2	Logistic Regression (Logit Link) . . . . .	453
12.8.3	Probit Regression (Probit Link) . . . . .	455
12.9	Generalized Linear Models . . . . .	456
12.9.1	Parameter Estimation . . . . .	460
12.9.2	Fisher Information, Tests and Confidence Intervals . . . .	461
12.10	Poisson Regression . . . . .	462
12.11	Overdispersion . . . . .	465
<b>A</b>	<b>Greek Letters</b>	<b>469</b>



<b>B</b>	<b>Summary of Brand-Name Distributions</b>	<b>471</b>
B.1	Discrete Distributions . . . . .	471
B.1.1	The Discrete Uniform Distribution . . . . .	471
B.1.2	The Binomial Distribution . . . . .	471
B.1.3	The Geometric Distribution, Type II . . . . .	472
B.1.4	The Poisson Distribution . . . . .	473
B.1.5	The Bernoulli Distribution . . . . .	473
B.1.6	The Negative Binomial Distribution, Type I . . . . .	473
B.1.7	The Negative Binomial Distribution, Type II . . . . .	474
B.1.8	The Geometric Distribution, Type I . . . . .	474
B.2	Continuous Distributions . . . . .	475
B.2.1	The Uniform Distribution . . . . .	475
B.2.2	The Exponential Distribution . . . . .	476
B.2.3	The Gamma Distribution . . . . .	476
B.2.4	The Beta Distribution . . . . .	477
B.2.5	The Normal Distribution . . . . .	477
B.2.6	The Chi-Square Distribution . . . . .	477
B.2.7	The Cauchy Distribution . . . . .	478
B.2.8	Student's $t$ Distribution . . . . .	478
B.2.9	Snedecor's $F$ Distribution . . . . .	479
B.3	Special Functions . . . . .	480
B.3.1	The Gamma Function . . . . .	480
B.3.2	The Beta Function . . . . .	480
B.4	Discrete Multivariate Distributions . . . . .	480
B.4.1	The Multinomial Distribution . . . . .	480
B.5	Continuous Multivariate Distributions . . . . .	482
B.5.1	The Uniform Distribution . . . . .	482
B.5.2	The Standard Normal Distribution . . . . .	482
B.5.3	The Multivariate Normal Distribution . . . . .	482
B.5.4	The Bivariate Normal Distribution . . . . .	484
<b>C</b>	<b>Addition Rules for Distributions</b>	<b>485</b>
<b>D</b>	<b>Relations Among Brand Name Distributions</b>	<b>487</b>
D.1	Special Cases . . . . .	487
D.2	Relations Involving Bernoulli Sequences . . . . .	487
D.3	Relations Involving Poisson Processes . . . . .	488
D.4	Normal, Chi-Square, $t$ , and $F$ . . . . .	488
D.4.1	Definition of Chi-Square . . . . .	488
D.4.2	Definition of $t$ . . . . .	488
D.4.3	Definition of $F$ . . . . .	489
D.4.4	$t$ as a Special Case of $F$ . . . . .	489

<b>E Eigenvalues and Eigenvectors</b>	<b>491</b>
E.1 Orthogonal and Orthonormal Vectors . . . . .	491
E.2 Eigenvalues and Eigenvectors . . . . .	493
E.3 Positive Definite Matrices . . . . .	496
<b>F Normal Approximations for Distributions</b>	<b>499</b>
F.1 Binomial Distribution . . . . .	499
F.2 Negative Binomial Distribution . . . . .	499
F.3 Poisson Distribution . . . . .	499
F.4 Gamma Distribution . . . . .	499
F.5 Chi-Square Distribution . . . . .	499
<b>G Maximization of Functions</b>	<b>501</b>
G.1 Functions of One Variable . . . . .	501
G.2 Concave Functions of One Variable . . . . .	503
G.3 Functions of Several Variables . . . . .	504
G.4 Concave Functions of Several Variables . . . . .	508
<b>H Projections and Chi-Squares</b>	<b>509</b>
H.1 Orthogonal Projections . . . . .	509
H.2 Chi-Squares . . . . .	511