0.1 Wrap Up

The topics in these notes are listed in both the table of contents and the index. This section briefly summarize the material, while highlighting some of the connections among the topics.

The probability axioms allow for a mathematical basis for modeling real-world systems with uncertainty, and allow for both discrete-type and continuous-type random variables. Counting problems naturally arise for calculating probabilities when all outcomes are equally likely, and a recurring idea for counting the total number of ways something can be done is to do it sequentially, such as in, "for each choice of the first ball, there are n_2 ways to choose the second ball," and so on. Working with basic probabilites includes working with Karnaugh maps, de Morgan's laws, and definitions of conditional probabilities and mutual independence of two or more events. Our intuition can be strengthened and many calculations made by appealing to the law of total probability and the definition of conditional probability. In particular, we sometimes look back, conditioning on what happened at the end of some scenario, and ask what is the conditional probability that the observation happened in a particular way–using Bayes rule. Binomial coefficients form a link between counting and the binomial probability distribution.

A small number of key discrete-type and continuous-type distributions arise again and again in applications. Knowing the form of the CDFs, pdfs, or pmfs, and formulas for the means and variances, and why each distribution arises frequently in nature and applications, can thus lead to efficient modeling and problem solving. There are relationships among the key distributions. For example, the binomial distribution generalizes the Bernoulli, and the Poisson distribution is the large n, small p limit of the Bernoulli distribution with $np = \lambda$. The exponential distribution is the continuous time version of the geometric distribution; both are memoryless. The exponential distribution is the limit of scaled geometric distributions, and the Gaussian (or normal) distribution, by the central limit theorem, is the limit of standardized sums of large numbers of independent, identically distributed random variables.

The following important concepts apply to both discrete-type random variables and continuous-type random variables:

- Independence of random variables
- Marginals and conditionals
- Functions of one or more random variables, the two or three step procedure to find their distributions, and LOTUS to find their means
- $E[X]$, $\mathrm{Var}(X)$, $E[XY]$, $\mathrm{Cov}(X,Y)$, $\rho_{X,Y}$, σ_X, and relationships among these.
- Binary hypothesis testing (ML rule, MAP rule as likelihood ratio tests)
- Maximum likelihood parameter estimation
- The minimum mean square error (MSE) estimators: δ^*, L^*, g^*
- Markov, Chebychev, and Schwarz inequalities (The Chebychev inequality can be used for confidence intervals; the Schwarz inequality implies correlation coefficients are between one and minus one.)
- Law of large numbers and central limit theorem

Poisson random processes arise as limits of scaled Bernoulli random processes. Discussion of these processes together entails the Bernoulli, binomial, geometric, negative geometric, exponential, Poisson, and gamma distributions.

Reliability in these notes is discussed largely in discrete settings–such as the outage probability for an $s-t$ network. Failure rate functions for random variables are discussed for continuous-time positive random variables only, but could be formulated for discrete time.

There are two complementary approaches for dealing with multiple random variables in statistical modeling and analysis, described briefly by the following two lists:

Distribution approach	Moment approach
joint pmf or joint pdf or joint CDF	means, (co)variances, correlation coefficients
marginals, conditionals	
independent	uncorrelated (i.e. $\mathrm{Cov}(X, Y) = 0$ or $\rho_{X,Y} = 0$)
g^*	L^*

That is, on one hand, it sometimes makes sense to postulate or estimate joint distributions. On the other hand, it sometimes makes sense to postulate or estimate joint moments, without explicitly estimating distributions. For jointly Gaussian random variables, the two approaches are equivalent. That is, working with the moments is equivalent to working with the distributions themselves. Independence, which in general is stronger than being uncorrelated, is equivalent to being uncorrelated for the case of jointly Gaussian random variables.