Stat 134: Final Review Sheet

Adam Lucas

Spring 2019

This review sheet is meant to provide a high level overview of concepts, techniques, and methods that we have covered throughout the course. There is less emphasis on computational practice (which you will have plenty of with the topic-specific review sheets) and more emphasis on synthesizing ideas to be prepared for the final. Note this is not an exhaustive review; it is intended to be a diagnostic tool for the start of your studies, or as a recap to be used at the end of your studies.

Fundamental Techniques

These common techniques seen throughout the course should be rock solid. If any of these results cause hesitation, now is a good time to return your notes where they were introduced and work through examples where they might be used.

a. *Rules of probability:* Find useful equalities for the following expressions:

  (i) $P(A|B)$;

  (ii) $P(\{A \cup B\}^c)$;

  (iii) $P(\{A \cap B\}^c)$.

b. *Constant and variable parts of densities:* Let X have density $f_X(x) = cx^3e^{-6x}$, where c is a constant. What is the distribution of X? Use this to evaluate $\int_0^\infty 5x^3e^{-6x}dx$.

c. *Independence and dependence:* How do we show that two events are independent? That two variables are independent? What is an easy way to show that two variables are dependent?

d. *Working with densities:* Remind yourself of the general approach for each of the following situations:

  (i) Obtaining F_X from f_X, and vice versa;

  (ii) Obtaining $f_X(x)$ from $f_{X,Y}(x,y)$;

  (iii) Obtaining $f_{X|Y}(x|y)$ from $f_{X,Y}(x,y)$ and the marginal densities f_X, f_Y.

Relationships Between Distributions

We have seen frequently how recognizing the transformations of random variables into other known distributions greatly simplifies problems. In the context of continuous RV problems, the variable at hand is often found by one of these transformations. For each of the following below, identify the distribution that results.

a. cX, where $X \sim \text{Exp}(\lambda)$, $c > 0$. We have proved this result using three different methods; recall these arguments.

b. Consider $X = U_{(3)}$ and $Y = U_{(7)}$, the 3rd and 7th order statistics of 10 iid Unif $(0, 1)$ RVs. What is the distribution of X/Y?

c. $X^2 + Y^2$, where X, Y are i.i.d. standard Normal. What is $\sqrt{X^2 + Y^2}$?

d. $2X + 3Y$, where $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. What if X, Y are bivariate normal with correlation $\rho = 0.6$?

e. Consider each of the following common discrete distributions: Poisson, Binomial, Geometric, and Hypergeometric. For which of these is the sum of two independent RVs a known distribution? Under what conditions?

Symmetry

A common argument used in the course is the notion of symmetry; this is sometimes an abstract concept but a very useful tool. Here are some questions related to this topic to help make it more concrete.

a. Under some conditions, we can quickly recognize the expectation of a random variable X to be zero. What are they?

b. Find the probability that the last ace in a standard, well-shuffled deck is at position 47 or greater.

c. You and I each roll a fair n sided die. Without using a summation, find the probability your roll is strictly greater than mine.

d. Let X, Y be independent $\mathcal{N}(0, \sigma^2)$ random variables. Without using Φ, find $P(X + 2Y > 0, X > 0)$. What is the reason behind this answer?

e. Let $X_1, X_2, \ldots, X_n$ be identically distributed, exchangeable variables such that their sum is always a constant. (E.g., the number of aces players 1 through n receive from a standard deck that is distributed to all players.) Find $Corr(X_i, X_j)$ for $i \neq j$.

To Infinity, and Beyond

Many distributions and problems in this class involve possible values up to $+\infty$. As we do not accept infinite sums, it is worth reviewing techniques for simplifying these problems.

a. Let $X \sim$ Geom (p) on $\{1, 2, \ldots\}$. It is not easy to directly find $\mathbb{E}(X)$ using the formula $\sum_{x=1}^{\infty} xP(X = x)$. We have shown three alternate methods for finding this expectation; what are they?

b. Two players simultaneously toss coins which land heads with probabilities p_1 and p_2 respectively. They continue until exactly one player's coin lands heads; that player is the winner. Show that the probability Player 1 wins is $\frac{p_1 q_2}{p_1 q_2 + q_1 p_2}$.

c. Find $\mathbb{E}(X(X-1))$ for $X \sim$ Pois (μ). Use this result to prove $Var(X) = \mu$.

d. Let Y have density $f_Y(y) = \frac{\lambda}{2} e^{-\lambda|y|}$, for $y \in \mathbb{R}$. With little computation, find $\mathbb{E}(|Y|)$ and $\mathbb{E}(Y)$.

Could You Rephrase That?

Often we can simplify probability expressions that may be difficult to directly compute by rephrasing them in terms of equivalent events, or using complements.

a. Suppose there are $3n$ people in a room, divided into groups of 3. What is the chance that there is at least one group where two members share the same birthday?

b. Continued from (a): Approximate this chance for large n.

c. Let $X \sim$ Gamma (r, λ), where r is a positive integer. Use what we know about the Poisson process to obtain the CDF of X.

d. Let X be an arbitrary random variable with an invertible CDF F_X. What is the random variable formed by $F_X(X)$?

Some Useful Results

The following are some results we have previously demonstrated in class which you may take for granted on the final (with some statement/justification of what you are using), but may not be on the reference sheet. Work to prove these results on your own if you can; if you are stuck, the proofs should be in your lecture notes.

a. Let $X \sim$ Pois (μ), $Y \sim$ Pois (λ), independent of each other. What is the distribution of $X + Y$? (Hint: use the binomial theorem, $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.)

b. Use Markov's Inequality to derive Chebyshev's Inequality.

c. Let $X \sim$ Exp (λ_X), $Y \sim$ Exp (λ_Y). What is $P(X < Y)$?

d. Find the distribution of $\min\{X_1, X_2, \ldots, X_n\}$, where the X_i's are independent Exp (λ) variables.

e. Suppose cars and trucks arrive at a bridge according to independent Poisson processes with rates λ_c and λ_r per minute respectively. Given that n vehicles arrive in t minutes, what is the distribution of $N_{c,t}$, the number of cars to arrive by time t?

Final Words

Do note that there are many subjects this worksheet has not addressed. A lot are computational, while some are more abstract. Of the latter, these include: Bayes' Rule, indicators, methods for calculating expectations, variances, and covariances, counting problems (hypergeometrics), and many others. It would be impossible to distill every technique in this course into a single final, much less a final review sheet.

Hopefully, however, this sheet gives you some tools for problem solving that may have been hinted at throughout the semester, but never gathered into one place. This kind of synthesis of materials and conceptual review, if taken seriously and viewed as a study aid rather than a worksheet, will go a long way to help you with reinforcing the intuition needed to approach problems on the final and in future classes.

Happy studying, and good luck with finals!

– The Stat 134 Team