Probability Theory: Statistics 530
Final Take Home Exam (Steele: Fall 2014)

INSTRUCTIONS. You may consult any books or articles that you find useful. If you use a result that is not from our text, attach a copy of the relevant pages from your source. You may use any software, including the internet, Mathematica, Maple, R, S-Plus, MatLab, etc. Attach any Mathematica (or similar) code that you use.

- You may **NOT** consult with any other person about these problems. If you have a question, even one that is just about the meaning of a question, please contact me directly rather than consult with a fellow student or other person. Any collaboration of any kind on any of these problems is a serious violation of the University Policy on Academic Integrity.
- I may post "bug reports" or clarifications on our web page, and you should regularly check for these on our web page.
- You should strive to make your answers as clear and complete as possible, but you can omit details that would be omitted by a professional mathematician (e.g. steps in integration by parts, Taylor expansions, known integrals, etc.) You need not be pedantic, but don't skip steps that an informed reader can't follow. If I don't know how to go from line $n$ to line $n+1$ something is missing. If you use Mathematica or a fact from a table, please say so **and document it.** Otherwise, I stare and stare at line $n$ wondering how you got to $n+1$ in your head while I can't. As a courtesy, never say "obviously;" If an assertion is obvious, it's obvious and there is no reason to say so.
- If you understand that your argument is incomplete or only heuristic, it may be appropriate to report what you have found, but such comments should be properly labeled as incomplete or heuristic.
- Use anything from anyplace, but do not steal. If you make use of an argument from some source, give credit to the source. If you find the complete (and correct!) solution to a problem in a book or on the internet, just print out the pages and attach them. You will get full credit. I intend to create original problems, but sometimes one gets "scooped."
- I will provide you with the Latex source for the problems. Please follow the format of giving your solution after my statement of the problem. Please do not renumber the problems. If there is a problem you do not do, just write "Problem not done" where you would have placed the solution.
- **As discussed in class, you MUST use and complete the cover page given at the website. Self-evaluation is hugely valuable.**

General Advice

1. In your solutions, please do not just write down things that you think are relevant when they do not add up to an honest solution. Such lists are useful to you when you are working on a problem, but if you offer such a list as a solution you are keeping yourself from having the “missing” idea.

2. If you can explain clearly something that you tried that did not work, this sometimes is worth a few points. Please do not abuse this offer. With experience, one learns that many sensible ideas do not work. Almost by definition, this is what separates the trivial from the non-trivial.

3. Try to keep in mind that a good problem requires that one “overcome some objection.” What distinguishes a problem from an exercise is that in a good problem a routine plan does not work. The whole point is to go past the place where routine ideas take you. Still, don’t shy away from the obvious; almost all of the “problems” here are “exercises.”

4. The justification of a formal computation is rarely hard, but sometimes it is the whole banana. The problems offered here are not about pedantry; they are about structure. Deal with rigor after you have a good plan.

5. If you do something **extra** that is valid, you can get “bonus” points. These special rewards cannot be determined in advance. They are usual small, but they can be substantial — and they do add up. In some past years, some students have gotten as much as 120%, i.e. all the basic points and 20% bonus points. Don’t make yourself nuts, but this is what is possible. The most common source of bonus points is for saying something particularly well. Clear, well-organize, solutions are gems. They deserve to be acknowledged.

Big Picture and Logistics

These instructions may seem overly detailed to you. It is true that they are detailed, but they evolved case by case. Each rule deals with some previous misunderstanding or missed opportunity.

**Due Date and Place:** A hard copy of this exam **with its completed self-evaluation cover sheet** is due in my mail box in JMHH Suite 400 on in my office JMHH 447 on the date given on our web page. You should *also email* me your solutions and Self-evaluation form as PDFs. You should label the files respectively:

- Solutions-530F14-Your-Name

- Self-Evaluation-530F14-Your-Name.

## The Problems

Problem 1. Let $\{X_n : n = 1, 2, \ldots\}$ be a sequence of integrable random variables. Let $\{\phi_n : n = 1, 2, \ldots\}$ be the corresponding characteristic functions. Suppose that we have

$$|1 - \phi_n(t)| \le A|t|^{1+\delta} \quad \text{for all } t \in \mathbb{R} \tag{1}$$

for some real values $A$ and $\delta > 0$ that do not depend on $n$.

1. Do we have $EX_n = 0$ for all $n$? Give a proof or a counter-example. Note: This part is an easy warm-up to make sure that you understand all of the assumptions in the problem.
2. Is this family of random variables $\{X_n : n = 1, 2, \ldots\}$ tight? Give a proof or a counter-example.
3. Is the family of random variables $\{X_n : n = 1, 2, \ldots\}$ dominated by an integrable random variable? Give a proof or a counter-example.
4. Is the family of random variables $\{X_n : n = 1, 2, \ldots\}$ uniformly integrable? Give a proof or a counter-example.

Problem 2. Let $\psi(t)$ be a characteristic function with

$$\psi'(0) = 0 \quad \text{and} \quad \psi''(0) = -1.$$

We then set $\phi_1(t) = \psi(t)$ and for $k = 1, 2, \ldots$ we set

$$\phi_{k+1}(t) = \frac{1}{2}\left\{\phi_k(t) + \phi_k(t)\psi(t)\right\}.$$

- Explain why $\phi_k(t)$ is a characteristic function for all $k = 1, 2, \ldots$. You can do this by describing an experiment that produces the corresponding random variable.
- Find the mean and variance of a random variable $S_k$ with characteristic function $\phi_k(t)$.
- Find the limit of the sequence $\phi_k(t/\sqrt{k})$ and give a probabilistic interpretation of the result. You can save work and show insight by using results we have proved rather than by repeating computations.
- State at least one generalization of this problem and sketch the argument for your generalization, or generalizations.

PROBLEM 3. The purpose of this problem is to examine an analog of the Levy theory for characteristic functions. For each $a \in \mathbb{R}$ we let $H_a(x) = \max(0, x-a)$ so $H_a$ is zero to the left of $a$ and from $a$ forward it grows linearly.

1. Prove that if $X$ and $Y$ are random variables with finite first moments, and if
$$EH_a(X) = EH_a(Y) \quad \text{for all } a \in \mathbb{R},$$
then $X$ and $Y$ are equal in distribution.

2. Let $Y$ and $X_n$, $n = 1, 2, \ldots$ be random variables such that $|Y| \leq Z$ and $\sup_n |X_n| \leq Z$ where $EZ < \infty$. Show that if
$$\lim_{n\to\infty} EH_a(X_n) = EH_a(Y) \quad \text{for all } a \in \mathbb{R},$$
then $X_n$ converges in distribution to $Y$.

3. Can you relax the hypothesis of the previous part? Address both the negative and the positive aspects of this question. For example, what if we just assume that $Y$ and $X_n$, $n = 1, 2, \ldots$ have first moments? What if we suppose these moments are uniformly bounded but that we do not have the domination condition of the previous part? What if we just assume that all of the random variables are integrable and that the sequence $X_n$, $n = 1, 2, \ldots$ is tight? In all of these instances give a proof or a counter example. You can expand on this problem for possible extra credit; the expansions are up to you and their formulation is part of the challenge.

PROBLEM 4. Suppose that $X_k$, $k = 1, 2, \ldots$ is a sequence of i.i.d. random variables with mean 0 and variance 1. Let $S_k = X_1 + X_2 + \cdots + X_k$ and let
$$h(\lambda) = \limsup_{n\to\infty} P(\max_{1\leq k\leq n} |S_k| \geq \lambda n^{1/2}).$$
Is it true that for each $p > 0$ there is a constant $C_p$ such that for all $\lambda > 0$ one has the inequality
$$h(\lambda) \leq C_p \lambda^{-p}.$$
Give a proof or a counter-example. Note: There is a book somewhere on the planet in some language that makes hash of this problem. If you find that book, you'll need to give the citation, sort out the hash, and explain your changes very carefully. My guess is that you'll save time by just trusting your own analysis. The problem is not hard, but it does require some facts and some clear thinking. You should not jump on the first idea you have; it could turn out to be a messy one. Let this problem ripen at least a little.

Problem 5. Let $X$ be a random variable and let $\phi$ be its characteristic function. Let $A$ be a nonnegative constant and consider the following inequality

$$|\phi(t) - \phi(s)| \le \sqrt{A|1 - \phi(t - s)|}. \tag{2}$$

Note that (2) implies that $\phi$ is uniformly continuous. This is something that we know always to be true, so the inequality (2) passes one test of feasibility. The inequality is also true when $s = t$, so it passes a second test. Further, one can see from (2) that if $\phi$ is equal to 1 in a neighborhood of zero, then $\phi$ is equal to 1 for all $t$. This is something that we proved in class by other means, so the inequality (2) passes yet another test. Still, (2) passes these tests for any value of $A$, so they are not very powerful tests. When we put on our pessimists hats, we can see pretty easily that (2) can fail if $A < 1$.

1. Show that if (2) holds for all random variables then $A \ge 2$. Specifically, show that for any $A < 2$ there is an $X$ for which (2) fails.

2. Is there an $A$, such that (2) holds for all $X$? Give a proof or a counterexample.

Problem 6. Let $f(x)$ be a probability density and assume that $f(x) = f(-x)$ for all $x$. Let $\phi(t)$ be the corresponding characteristic function and assume that $\phi(t) > 0$ for all $t$. Note, for example, the standard normal has these properties. Now, for $a \ne 0$ we consider the ratio

$$\phi_a(t) = \frac{2\phi(t) - \phi(t+a) - \phi(t-a)}{2(1 - \phi(a))}.$$

1. Confirm that $\phi_a(t)$ is a characteristic function by finding the density $f_a(x)$ that has characteristic function $\phi_a(t)$. The formula for $f_a(x)$ can be written as $f(x)m(a,x)$ where $m(a,x)$ is a function that you can write rather simply in terms of $\phi$ and trigonometric functions.

2. Look up a good version of the Riemann-Lebesgue lemma, (the version in Exercise 1.4.4 of Durrett is almost good enough; the suggested proof *is good enough*). Use this to show that $\phi_a(t) \to \phi(t)$ as $a \to \infty$.

3. Show that $f_a(x)$ does not converge to $f(x)$ as $a \to \infty$. Given part 2, does this contradict Levy's continuity theorem or Levy's inversion formula? Explain your reasoning.

4. For extra credit, you can comment on $\phi_a(t)$ as $a \to 0$. Is the limit an honest characteristic function? What if you assume that the density of $f$ corresponds to a random variable with a finite second moment?

Problem 7. Let $\phi(t)$ be the characteristic function of the random variable $X$ that may or may not have have any atoms.

1. Fill in the details of the hints given in Durrett's problem 3.3.7 to prove that
$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |\phi(t)|^2\,dt = \sum_{x\in\mathbb{R}} P(X=x)^2.$$
Note that the last sum has uncountably many summands but at most countably many of these are nonzero.

   This is an easy problem, so do a good job of exposition — don't make the poor reader bounce all over the book like Rick does. In fact, you should aim for a write up that is reaonably self-contained.

2. Check this relationship by giving an independent proof that it holds for $X$ that takes each of the values 1 and $-1$ with probability $1/2$.

3. Next assume that $X$ has no atoms, so the the sum given above is zero. Now assume that $|\phi'(t)| \le B$ for some $B < \infty$ and all $t$. Show that $\phi(t) \to 0$ as $t \to \infty$. Note that we could get this conclusion trivially from the Riemann-Lebesgue lemma if we were to assume that $X$ has a density. Here we assume that $X$ is non-atomic, but this is much weaker than assuming that $X$ has a density.

Problem 8. Assume that $X_n$, $n = 1, 2, \dots$ is a sequence of independent random variables with mean zero and finite fourth moments. More precisely, assume $EX_n^4 < \infty$ for each $n$, but do not assume that there is a uniform bound on these moments; *a priori*, they may be arbitrarily large. Consider the two conditions:

$$EX_n^4 \le 8(EX_n^2)^2 \quad \text{for all } n = 1, 2, \dots \tag{3}$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\left(\sum_{k=1}^{n} X_k^2\right)^2 < \infty \quad \text{with probability 1.} \tag{4}$$

Assume condition (3) and (4), set $S_n = X_1 + X_2 + \cdots + X_n$ and prove
$$P(\lim_{n\to\infty} S_n/n = 0) = 1.$$

Bonus Points: Can you drop condition (4) and get the same conclusion? Give a proof or a counter-example. If you can't drop condition (4) perhaps you can weaken it for more bonus points.

Problem 9. You play a coin flipping game with a friend, and you keep track of your net winnings. Naturally, your net winnings start at zero. The odds are fair, since you are playing with a friend. You follow a natural strategy: (1) when your net winning is zero or positive you bet one dollar and (2) when your net winning is negative you bet two dollars.

1. What is the probability that you win 100 dollars before you lose 301 dollars?
2. How long does the game last?
3. Generalize these results in any way that pleases you. An especially nice generalization will get bonus points, but everyone needs to come up with some generalization. You can start as modestly as you like, and it is best not to begin with a big jump.

Problem 10. For the following assertions give a proof or a counter-example

- There is a martingale $\{M_n, \mathcal{F}_n\}$ such that $n^{-1}M_n$ goes to infinity with probability one as $n \to \infty$.
- There is a martingale $\{M_n, \mathcal{F}_n\}$ such that with probability one:

$$\liminf_{n\to\infty} M_n = -\infty, \quad \limsup_{n\to\infty} M_n = \infty, \quad \text{and} \sum_{n=1}^{\infty} \mathbb{1}(M_n = 0) = \infty.$$

Problem 11. Suppose that $X_n$ and $Y_n$ are random variables that are adapted to the filtration $\{\mathcal{F}_n\}$. Suppose that for all $n$ we have $EX_n^2 < \infty$ and $EY_n^2 < \infty$ but these expectations are not necessarily uniformly bounded. Suppose also that we have

$$E(X_{n+1}^2 \,|\, \mathcal{F}_n) \le X_n^2 + Y_n^2 \quad \text{and} \quad \sum_{n=1}^{\infty} Y_n^2 < \infty \, a.s.$$

1. Is it true that $X_n^2$ converges as $n \to \infty$? Give a proof or a counter-example.
2. Is it true that $X_n$ converges as $n \to \infty$? Give a proof or a counter-example.

Problem 12.
Let $X_j$, $j = 1, 2, ...$ be i.i.d uniform on $[0,1]$ and let

$$Y_j(n) = \frac{X_j}{X_1 + X_2 + \cdots + X_n} \quad 1 \le j \le n.$$

Note that the $Y_j$ are *dependent*. This is a useful recipe for choosing a point $(Y_1, Y_2, \ldots, Y_n)$ at random out of the unit simplex in $\mathbb{R}^n$.

- Find the expectation of $Y_1(n)$ exactly.

- Show that $\operatorname{Var} Y_1(n) = c/n^2 + o(1/n^2)$ and determine the constant $c$. If you get something a little sharper — fine. For the moment this is all we should need.

- It is probably intuitive to you that if $n$ is large and $m$ is large but much smaller than $n$ then the partial sum $Y_1(n) + Y_2(n) + \cdots + Y_m(n)$ can be normalized in a way that gives one a CLT style approximation to its distribution. Formulate and prove some theorem of this kind.

  You do not need to press for the most general result you can get. You should focus on making sure that what you say is true and making sure that you can prove what you claim.

- For your proof, you should take a hint from the fact that for large $n$ the law of large numbers gives you some control over the size of the numerator that defines the $Y$'s. You can introduce a good set where

$$G = \{\omega : n(1/2 - \epsilon) \leq X_1 + X_2 + \cdots + X_n \leq n(1/2 + \epsilon)\}$$

  and then you can reason honestly using the good set $G$ and the corresponding bad set $B$.

  You have to use this information judiciously. If you use the word "conditional probability" here, you are almost certainly cutting yourself adrift on an iceberg. The notion is not needed here and it is very hard to handle appropriately. I'd try to use $G$ then use appropriate inequalities to "get rid" of $G$. Also, keep in mind that an identity is nothing but two inequalities where someone let something or (some things) go to a limit — you know, it's the old squeeze play — time and time again.