Math 561 — Spring 2006 — Test Solutions

Total points: 100. 75 minutes. Show ALL your working and make your explanations as full as possible. Electronic devices are not allowed; neither are books or notes.

1: (30 points) **Do part (a) or part (b), but not both.**

(a) Prove that if \( \mathcal{A}_1 \) and \( \mathcal{A}_2 \) are \( \pi \)-systems that are independent, then \( \sigma(\mathcal{A}_1) \) and \( \sigma(\mathcal{A}_2) \) are independent.

*Solution.* Durrett Theorem 1.4.2 (page 24), with \( n = 2 \).

Your solution should clearly point out where the independence hypothesis on \( \mathcal{A}_1 \) and \( \mathcal{A}_2 \) is used, namely to show that \( \mathcal{A}_1 \subset \mathcal{L} \).

(b) \((L^4 \text{ Strong Law})\) Let \( X_1, X_2, \ldots \) be i.i.d. with \( EX_i = \mu \) and \( EX_i^4 < \infty \). Write \( S_n = X_1 + \cdots + X_n \). Prove \( S_n/n \to \mu \) a.s.

*Solution.* Durrett Theorem 1.6.5 (page 48).

Your solution needs to clearly show where the independence of \( X_1, X_2, \ldots \) is used, namely to get that \( E(X_iX_jX_kX_\ell) = EX_iEX_jEX_kEX_\ell \) if \( i, j, k, \ell \) are all distinct, and so on.
2: (25=13+12 points) Let \( X_1, X_2, \ldots \) be i.i.d. standard normal random variables (mean 0 and variance 1). It is known by Theorem 1.1.4 on page 6 of Durrett that
\[
(6x)^{-1}e^{-x^2/2} \leq P(X_i \geq x) \leq (2x)^{-1}e^{-x^2/2} \quad \text{when } x > 2.
\]
(a) Prove \( \lim \sup_n (X_n/\sqrt{2 \log n}) \leq 1 \) a.s.

\[\text{Solution.} \quad \text{Let } \varepsilon > 0. \text{ Then}
\begin{align*}
\sum_{n \geq 3} P\left( \frac{X_n}{\sqrt{2 \log n}} \geq \sqrt{1 + \varepsilon} \right) &= \sum_{n \geq 3} P\left( X_n \geq \sqrt{2(1 + \varepsilon) \log n} \right) \\
&\leq \sum_{n \geq 3} \frac{1}{2\sqrt{2(1 + \varepsilon) \log n}} \exp\left( -2(1 + \varepsilon)(\log n)/2 \right) \\
&\leq \sum_{n \geq 3} n^{-(1+\varepsilon)} < \infty
\end{align*}\]
where we have used in the denominator that \( \log n > 1 \) when \( n \geq 3 \).

Borel–Cantelli I (Lemma 1.6.1 on page 46) now implies that
\[P\left( \frac{X_n}{\sqrt{2 \log n}} \geq \sqrt{1 + \varepsilon} \text{ i.o.} \right) = 0,
\]
so that \( P(\lim \sup_n (X_n/\sqrt{2 \log n}) > \sqrt{1 + \varepsilon}) = 0 \). Hence \( \lim \sup_n (X_n/\sqrt{2 \log n}) \leq \sqrt{1 + \varepsilon} \) a.s., and now letting \( \varepsilon \downarrow 0 \) through a discrete sequence of values shows that \( \lim \sup_n (X_n/\sqrt{2 \log n}) \leq 1 \) a.s.

Note we have not used independence of the \( X_i \), in Borel–Cantelli I.

(b) Prove \( \lim \sup_n (X_n/\sqrt{2 \log n}) \geq 1 \) a.s.

\[\text{Solution.} \quad \text{We don’t need to use } (1 - \varepsilon) \text{ in this proof, because we can simply show:}
\begin{align*}
\sum_{n \geq 2} P\left( \frac{X_n}{\sqrt{2 \log n}} \geq 1 \right) &= \sum_{n \geq 2} P\left( X_n \geq \sqrt{2 \log n} \right) \\
&\geq \sum_{n \geq 2} \frac{1}{6\sqrt{2 \log n}} \exp\left( -(2 \log n)/2 \right) \\
&= \frac{1}{6\sqrt{2}} \sum_{n \geq 2} \frac{1}{n^{\sqrt{2}}} \\
&\geq \frac{1}{6\sqrt{2}} \int_{2}^{\infty} \frac{1}{x^{\sqrt{2}}} \, dx = \infty.
\end{align*}\]

Borel–Cantelli II (Lemma 1.6.6 on page 49) now implies that
\[P\left( \frac{X_n}{\sqrt{2 \log n}} \geq 1 \text{ i.o.} \right) = 1,
\]
so that \( \lim \sup_n (X_n/\sqrt{2 \log n}) \geq 1 \) a.s.

Note we relied on independence of the \( X_i \), when we used Borel–Cantelli II.
(c) Extra credit, 10 points. Prove $\lim sup_n \frac{S_n}{\sqrt{2n \log n}} \leq 1$ a.s., where $S_n = X_1 + \cdots + X_n$.

**Warning:** part (c) originally had a typo, stating that “... = 1 a.s.” Here we correctly prove “... ≤ 1 a.s.”

**Solution.** $S_n$ is a sum of independent normal random variables, and hence is itself a normal random variable, with mean $E S_n = EX_1 + \cdots + EX_n = 0$, and variance $\text{Var}(S_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) = n$ (using independence of the $X_i$). Thus $S_n/\sqrt{n}$ is normal with mean 0 and variance 1.

Now part (a) above gives $\lim sup_n \frac{S_n/\sqrt{n}}{\sqrt{2 \log n}} \leq 1$ a.s. [Note we cannot use part (b) because the $S_n$ are dependent; exercise!]

**Remark.** A more precise result than part (c) is the **Law of the Iterated Logarithm** (Theorem 7.9.7 on page 434):

$$\lim sup_n \frac{S_n}{\sqrt{2n \log \log n}} = 1 \text{ a.s.}$$

**So what does Problem 2 tell us?**

(a) and (b) together say that if you repeatedly take independent measurements of a quantity that has a standard normal distribution, then in the long run you expect to get arbitrarily large measurements. More precisely, after about $n$ measurements you expect to observe some measurements as large as $\sqrt{2 \log n}$.

This might seem surprising, since the probability of a standard normal random variable taking a value even greater than 3 is only 0.0014. But maybe it is not so amazing after all, because when $n = 10^6$ we calculate $\sqrt{2 \log n} \approx 5.7$.

As applied mathematicians like to say, the logarithm is almost a constant function.

For part (c), consider a random walk on $\mathbb{R}$ that starts at the origin, with the size of each step chosen from the standard normal distribution. Then $S_n$ gives the position after $n$ steps. Part (c) says that in the long run, you expect the random walk to get distance at most $\sqrt{2n \log n}$ away from the origin after $n$ steps.
3: (25=15+10 points) Let \( f : [0, 1] \rightarrow \mathbb{R} \) be Borel measurable with \( \int_0^1 f(u)^2 \, du < \infty \). Let \( U_1, U_2, \ldots \) be i.i.d. uniform random variables on \([0, 1]\). Let

\[
I_n = \frac{f(U_1) + \cdots + f(U_n)}{n}.
\]

(a) Show \( I_n \rightarrow I := \int_0^1 f(u) \, du \) a.s.

**Solution.** \( f(U_i) \) is an \( L^1 \) random variable (in fact, \( L^2 \)), with

\[
E|f(U_i)| = \int_0^1 |f(u)| \, du \quad \text{since } U_i \text{ is uniform on } [0, 1]
\]

\[
\leq \int_0^1 f(u)^2 \, du \quad \text{by Jensen’s inequality}
\]

\[
< \infty.
\]

Hence the expectation \( E f(U_i) = \int_0^1 f(u) \, du = I \) is well defined.

And \( f(U_1), f(U_2), \ldots \) are independent, because functions of independent random variables are independent (Corollary 1.4.5), and they are also identically distributed of course. Therefore

\[
I_n = \frac{f(U_1) + \cdots + f(U_n)}{n} \rightarrow E f(U_i) = I \quad \text{a.s.}
\]

by the \( L^1 \) Strong Law of Large Numbers (Theorem 1.7.1 on page 55).

(b) Let \( a > 0 \). Use Chebyshev’s inequality to show

\[
P(|I_n - I| > an^{-1/2}) \leq a^{-2} \text{Var}(f(U_i)).
\]

**Solution.**

\[
P(|I_n - I| > an^{-1/2}) \leq \frac{(an^{-1/2})^{-2} E(I_n - I)^2}{a^{-2} n \text{Var}(I_n)} = a^{-2} n \text{Var}(I_n) = a^{-2} n^{-1} \text{Var}(n I_n)
\]

\[
= a^{-2} n^{-1} \text{Var}\left( \sum_{m=1}^n f(U_m) \right)
\]

\[
= a^{-2} n^{-1} \sum_{m=1}^n \text{Var}(f(U_m)) \quad \text{using independence}
\]

\[
= a^{-2} \sum_{m=1}^n \text{Var}(f(U_m)) \quad \text{using identical distribution.}
\]

Incidentally, observe that \( \text{Var}(f(U_i)) \) is finite because \( f(u) \) is in \( L^2 \).
Do problem 3 or problem 4, but not both.

4: (25+10+10+5 points) Let $Z_n$ be a Poisson random variable with parameter $\lambda = n$.

(a) Show $Z_n/n \to 1$ a.s.

Solution. Decompose $Z_n = X_1 + \cdots + X_n$ as a sum of Poisson i.i.d. random variables, with each $X_i$ having parameter $\lambda = 1$. (Recall that a sum of independent Poisson random variables is again a Poisson random variable.) Then $EX_i = 1$ and so $Z_n/n \to 1$ a.s. by the Strong Law of Large Numbers. (The $L^4$ Strong Law and the $L^1$ Strong Law both apply.)

(b) Roughly sketch a normal probability density that approximates the distribution for $Z_n$. Indicate the approximate location and width of your sketch.

Solution. $EX_i = 1$ and $\text{Var}(X_i) = 1$, and so the Central Limit Theorem 2.4.1 applied to $X_1, X_2, \ldots$ says that the distribution of $(Z_n - n)/\sqrt{n}$ is approximately standard normal, when $n$ is large. In other words, $Z_n$ is approximately normal with mean $n$ and variance $n$, standard deviation $\sqrt{n}$. This can be sketched roughly as:

(c) Customers arrive at a store in a Poisson fashion at a rate of 100 per hour, on average. Estimate the probability that more than 120 customers arrive during the next hour.

Solution. With $n = 100$ we have

$$P(Z_{100} > 120) = P\left(\frac{Z_{100} - 100}{\sqrt{100}} > \frac{120 - 100}{\sqrt{100}}\right)$$

$$\approx P(\chi > 2)$$

$$= 1 - 0.9772 = 0.0228$$

to 4 decimal places, using the Table for the standard normal distribution $\chi$. 

5
5: (20 points) Do part (a) or part (b), but not both.

(a) Let $X$ be a uniform random variable on the interval $[-1, 1]$. Compute the characteristic function $\varphi_X(t)$. Then show there cannot exist i.i.d. random variables $Y$ and $Z$ with $Z - Y = X$.

Solution.

$$\varphi_X(t) = E(e^{itX}) = \int_{-1}^{1} e^{itx} \frac{1}{2} dx = \frac{\sin t}{t}.$$ 

Now suppose $Y$ and $Z$ are i.i.d. Then

$$\varphi_{Z-Y}(t) = E(e^{itZ}e^{-itY}) = E(e^{itZ})E(e^{-itY}) \quad \text{by independence}$$

$$= E(e^{itZ})E(e^{itY})$$

$$= E(e^{itZ})E(e^{itZ}) \quad \text{by identical distribution}$$

$$= |\varphi_Z(t)|^2 \geq 0.$$ 

This nonnegative function cannot equal $\varphi_X(t) = (\sin t)/t$, because $\sin t$ is negative sometimes (for example at $t = 3\pi/2$).

(b) Prove that if $X_n \Rightarrow X$ and $c_n \to 0$ then $X_n - c_n \Rightarrow X$. (Use only the definition of weak convergence; do not use any theorems or exercises.)

Solution. Let $F(x) = P(X \leq x)$ be the distribution function of $X$. Suppose $x$ is a point of continuity of $F$. Choose $\varepsilon > 0$ such that $x + \varepsilon$ is also a point of continuity; note this works for all except countably many $\varepsilon$-values, because $F$ has at most countably many discontinuities. Then

$$P(X_n - c_n \leq x) = P(X_n \leq x + c_n)$$

$$\leq P(X_n \leq x + \varepsilon) \quad \text{for all large } n, \text{ since } c_n \to 0$$

$$\to P(X \leq x + \varepsilon) \quad \text{because } X_n \Rightarrow X.$$ 

That is, we have shown

$$\limsup_n P(X_n - c_n \leq x) \leq P(X \leq x + \varepsilon) = F(x + \varepsilon),$$

for all except countably many $\varepsilon$-values. Letting $\varepsilon \to 0$ implies

$$\limsup_n P(X_n - c_n \leq x) \leq P(X \leq x) = F(x),$$

using here that $F$ is continuous at $x$.

A similar argument gives that $P(X \leq x)$ is a lower bound on the lim inf, and so in fact equality holds and $\lim_n P(X_n - c_n \leq x) = P(X \leq x)$, which proves that $X_n - c_n \Rightarrow X$.

Remark. We have already used the result in part (b) several times in class; now you have a proof.