Statistics 522, Midterm Exam

Wellner; 2/17/2017

1. (24 points). **Define** *three* of the following five terms:
   (a) The conditional expectation of a random variable X defined on $(\Omega, \mathcal{A}, P)$ given a (sub-) sigma-field $\mathcal{D}$ of $\mathcal{A}$
   (b) A martingale, sub-martingale, and super-martingale.
   (c) A stopping time T (relative to a filtration $\mathcal{A}_n$).
   (d) The compensator of a sub-martingale.
   (e) The predictable variation process $\langle X \rangle_n$ corresponding to a martingale $\{X_n, \mathcal{A}_n\}$ with $E(X_n^2) < \infty$ for all n.

   Solution: See Course notes.

2. (27 points). Give careful **statements** of *three* of the following six theorems or results:
   (a) The three-series theorem.
   (b) The (reverse) martingale version of the strong law of large numbers.
   (c) Doob's maximal inequality for a sub-martingale.
   (d) Doob's decomposition theorem for a sub-martingale $\{X_n, \mathcal{A}_n\}$.
   (e) Doob's L_r maximal inequality for a sub-martingale $\{X_n, \mathcal{A}_n\}$;
   and Doob's L_r maximal inequality for a martingale $\{X_n, \mathcal{A}_n\}$.
   (f) Three properties of conditional expectation.

   Solution: See Course notes.

3. (30 points). Suppose that Y is a random variable on the probability space $(\Omega, \mathcal{A}, P)$ with $Y \in L_2(P)$. Let $\mathcal{D} = \sigma[X_1, X_2, \ldots]$ for some sequence of random variables $\{X_1, X_2, \ldots\}$.
   (a) Suppose that $Z \in L_2(P)$ is $\mathcal{D}-$ measurable. Show that $E\{Z(Y - E(Y|\mathcal{D}))\} = 0$.
   (b) Interpret this geometrically in the Hilbert space $H = L_2(P)$ with inner product $\langle U, V \rangle = E(UV)$. (Draw a picture!)
   (c) Show that $E(Y - Z_0)^2 \ge E[Y - E(Y|\mathcal{D})]^2$ for all $\mathcal{D}-$measurable random variables Z_0 with equality if and only if $Z_0 = E(Y|\mathcal{D})$ almost surely.

   Solution: (a) This follows since

$$\begin{aligned} E\{Z(Y - E(Y|\mathcal{D}))\} &= E\{ZY\} - E\{ZE(Y|\mathcal{D})\} \ \text{ by linearity of cond. expectation} \\ &= E\{ZY\} - E\{E(ZY|\mathcal{D})\} \ \text{ since } Z \text{ is } \mathcal{D}\text{- measurable} \\ &= E\{ZY\} - E\{ZY\} = 0. \end{aligned}$$

(b) The geometric interpretation is simply that the error in "prediction" of Y by $E(Y|\mathcal{D})$ given all the information in $\mathcal{D}$ is orthogonal to the subspace of $L_2(P)$ given by

all square-integrable function Z which are $\mathcal{D}-$measurable: $Y - E(Y|\mathcal{D}) \perp H_{\mathcal{D}}$ where $H_{\mathcal{D}} \equiv \{Z : \; Z \;$ is $\mathcal{D} -$ measurable, $\; E(Z^2) < \infty\}$.
(c) This follows easily from (a) since

$$\begin{aligned}
E(Y - Z_0)^2 &= E(Y - E(Y|\mathcal{D}) + E(Y|\mathcal{D}) - Z_0)^2 \\
&= E[(Y - E(Y|\mathcal{D}))^2] + 2E\{(Y - E(Y|\mathcal{D}))(E(Y|\mathcal{D}) - Z_0)\} + E[(E(Y|\mathcal{D}) - Z_0)^2] \\
&= E[(Y - E(Y|\mathcal{D}))^2] + 0 + E[(E(Y|\mathcal{D}) - Z_0)^2] \;\text{ by (a) with } Z \equiv E(Y|\mathcal{D}) - Z_0 \\
&\geq E[(Y - E(Y|\mathcal{D}))^2] + 0 + 0
\end{aligned}$$

with equality if and only if $Z_0 = E(Y|\mathcal{D})$ almost surely.

Do one of the following two problems:

4. (36 points) Suppose that $X_1, X_2, \ldots$ are i.i.d. with $E|X_1| < \infty$. Let $S_n \equiv X_1 + \cdots + X_n$ and $\mathcal{A}_n \equiv \sigma[X_1, \ldots, X_n]$. Let $\mu \equiv E(X_1)$.
(a) Show that $\{S_n - n\mu, \mathcal{A}_n\}_{n \geq 1}$ is a martingale.
(b) Let T be a stopping time with $E(T) < \infty$. Show that the random variable

$$Y \equiv \sum_{k=1}^{T} |X_k| = \sum_{k=1}^{\infty} |X_k| 1_{[T \geq k]}$$

has $E(Y) = E|X_1| \cdot E(T) < \infty$. *Hint: T is a stopping time; recall an expression for $E(T)$ in terms of the probabilities $P(T \geq k)$.*
(c) Use (a) and (b) to show that $E(S_T) = \mu E(T)$. (This is known as *Wald's equation.*)

Solution: (a) First, $E|S_n - n\mu| \leq \sum_{i=1}^n E|X_i| + n|\mu| < \infty$ for each $n \geq 1$. Furthermore,

$$\begin{aligned}
E(S_{n+1} - (n+1)\mu|\mathcal{A}_n) &= E\{X_{n+1} - \mu + S_n - n\mu|\mathcal{A}_n\} \\
&= E\{X_{n+1} - \mu\} + S_n - n\mu \;\text{ since } X_{n+1} \text{ is independent of } \mathcal{A}_n \\
&\qquad \text{and } \; S_n - n\mu \; \text{ is } \mathcal{A}_n - \text{measurable} \\
&= 0 + S_n - n\mu \;\text{ a.s.}
\end{aligned}$$

Thus $\{S_n - n\mu, \mathcal{A}_n\}_{n \geq 1}$ is a martingale.
(b) Note that $[T \geq k] = [T \leq k-1]^c \in \mathcal{A}_{k-1}$ and thus $1_{[T \geq k]}$ and X_k are independent. Thus

$$\begin{aligned}
E(Y) &= E\left\{\sum_{k=1}^{T} |X_k|\right\} = E\left\{\sum_{k=1}^{\infty} |X_k| 1_{[T \geq k]}\right\} \\
&= \sum_{k=1}^{\infty} E|X_k| \cdot P(T \geq k) = E|X_1| \sum_{k=1}^{\infty} P(T \geq k) = E|X_1| \cdot E(T)
\end{aligned}$$

where we used $E(T) = \sum_{k=1}^{\infty} P(T \geq k)$ in the last line.
(c) For any integer n, $T \wedge n$ is a bounded stopping time, and hence by the simple

optional sampling theorem it follows from (a) that $\{S_1 - \mu, S_{T\wedge n} - (T \wedge nZ)\mu\}$ is a two-term martingale: in particular

$$0 = E\{S_{T\wedge n} - (T \wedge n)\mu\}.$$

Thus we have

$$\mu E(T \wedge n) = E\{S_{T\wedge n}\} \tag{1}$$

for each $n \geq 1$. Letting $n \to \infty$, the left side converges to $\mu E(T)$ by the monotone convergence theorem. On the other hand $|S_{T\wedge n}| \leq Y$ where $E(Y) < \infty$ by (b). Thus the right side of (1) converges to $E(S_T)$ by the dominated convergence theorem. Thus we conclude that

$$E(S_T) = \mu E(T);$$

i.e. Wald's identity holds.

5. (36 points) Suppose that $X_1, X_2, \ldots$ are i.i.d. Rademacher random variables (so that $P(X_i = \pm 1) = 1/2$), and let $S_n = X_1 + \cdots + X_n$.
(a) Show that $\{S_n, \mathcal{A}_n\}$ and $\{S_n^2 - n, \mathcal{A}_n\}$ are martingales.
(b) Let $\tau \equiv \min\{n : \ S_n = 1\}$. In class we used an exponential martingale to show that $p \equiv P(\tau < \infty) = 1$. Show this again by conditioning on X_1 and proving that p satisfies $p = (1/2)(p^2 + 1)$, which has the unique solution $p = 1$ in $[0, 1]$. (You may proceed heuristically by assuming that the strong Markov property holds.)
(c) Now suppose that the X_n's and S_n's are as above, but now let $\tau \equiv \min\{n \geq 1 : \ S_n = b \text{ or } S_n = -a\}$ where a, b are positive integers. Show that $E(S_\tau) = 0$ and $E(S_\tau^2) = E\tau = ab$.

Solution: (a) Since the X_i's have $E(X_i) = 0$ and $Var(X_i) = 1$,

$$E(S_{n+1}|\mathcal{A}_n) = E(X_{n+1} + S_n|\mathcal{A}_n) = E(X_{n+1}) + S_n = S_n \ \text{ a.s.}$$

and

$$\begin{aligned} E\{(S_{n+1}^2 - (n+1)|\mathcal{A}_n\} &= E\{(X_{n+1} + S_n)^2 - (n+1)|\mathcal{A}_n\} \\ &= E\{X_{n+1}^2 + 2X_{n+1}S_n + S_n^2|\mathcal{A}_n\} - (n+1) \\ &= E(X_{n+1}^2) - 1 + 2S_n E(X_{n+1}) + S_n^2 - n \ \text{ a.s.} \\ &= 0 + 0 + S_n^2 - n = S_n^2 - n^2. \end{aligned}$$

Thus $\{S_n, \mathcal{A}_n\}$ and $\{S_n^2 - n, \mathcal{A}_n\}$ are both martingales.
(b) With $p = P(\tau < \infty)$ we have

$$\begin{aligned} p &= P(\tau < \infty) = E1_{[\tau<\infty]} = E\{E\{1_{[\tau<\infty]}|X_1\}\} \\ &= E\{1_{[X_1=1]} + p^2 1_{[X_1=-1]}\} = 1 \cdot (1/2) + p^2 \cdot (1/2) = (1/2)(1 + p^2). \end{aligned}$$

But we note that with $h(p) \equiv (1/2)(1+p^2)$ we have $h(p) \geq p$ for all $p \in [0,1]$ with equality if and only if $p = 1$ since $h(p) - p = (1/2)(p^2 - 2p + 1) = (1/2)(p-1)^2 \geq 0$ with equality if and only if $p = P(\tau < \infty) = 1$. The optional sampling theorem clearly fails for this τ and the martingale $\{S_n\}$ since $S_\tau = 1$ a.s. and hence $1 = E(S_\tau) > ES_1 = 0$.
(c) For the two-sided stopping time $\tau \equiv \tau_{ab}$, $\tau \wedge n$ is a bounded stopping time and hence the simple optional sampling theorem applied with the martingales $\{S_n\}$ and $\{S_n^2 - n\}$ yield

$$0 = E(S_{\tau\wedge n}) \qquad \text{and} \qquad 0 = E(S_{\tau\wedge n}^2 - \tau \wedge n).$$

In this case $|S_{\tau\wedge n}| \leq a \wedge b$, so the dominated convergence theorem, together with the monotone convergence theorem, yields

$$0 = ES_\tau \quad \text{and} \quad 0 = E(S_\tau^2 - \tau),$$

Since $S_\tau \in \{-a, b\}$ with probability 1, the first identity yields, with $p_b \equiv P(S_\tau = b)$,

$$0 = bp_b + (-a)(1 - p_b),$$

and solving this for p_b yields $p_b = a/(a+b)$. Thus the second identity yields

$$E(\tau) = E(S_\tau^2) = b^2 p_b + a^2(1 - p_b) = b^2 \frac{a}{a+b} + a^2 \frac{b}{a+b} = ab.$$

Do one of the following two problems: 6 or 7

6. (36 points). Suppose that $\xi_1, \xi_2, \ldots$ are i.i.d. Uniform$(0,1)$ random variables, and let $\mathbb{G}_n(t) = n^{-1}\sum_{i=1}^n 1_{[0,t]}(\xi_i)$ be the empirical distribution function of the ξ_i's. Define $\mathcal{F}_n(t) \equiv \sigma\{\mathbb{G}_n(s): \ s \le t \le 1\}$ for $0 < t \le 1$.
(a) Show that $\{\mathbb{G}_n(t)/t, \mathcal{F}_n(t): \ 0 < t \le 1\}$ is a reverse martingale in the following sense: if $0 < s < t < 1$, then

$$E\left\{\frac{\mathbb{G}_n(s)}{s}\Big|\mathcal{F}_n(t)\right\} =_{a.s.} \frac{\mathbb{G}_n(t)}{t}\,.$$

Hint: Note that $(n\mathbb{G}_n(s), n(\mathbb{G}_n(t)-\mathbb{G}_n(s)), n(1-\mathbb{G}_n(t)))$ has a multinomial distribution $\text{Mult}_3(n,(s,t-s,1-t))$, and hence the
conditional distribution of $(n\mathbb{G}_n(s), n(\mathbb{G}_n(t)-\mathbb{G}_n(s)))$ given $\mathcal{F}_n(t)$ is the same as the conditional distribution of $(n\mathbb{G}_n(s), n(\mathbb{G}_n(t)-\mathbb{G}_n(s)))$ given $n(1-\mathbb{G}_n(t))$, namely $\text{Mult}_2(n\mathbb{G}_n(t), (s/t, 1-s/t))$.
(b) Use the result of (a) to show that for $0 < a < 1$ and $\lambda > 0$

$$P(\sup_{a\le t\le 1}\frac{\mathbb{G}_n(t)}{t} > \lambda) \le \exp(-nah(\lambda))\,.$$

where $h(\lambda) = \lambda(\log\lambda - 1) + 1$.
(c) use the result of (c) to show that if $a = a_n$ sastisfies $na_n \to \infty$, then

$$\sup_{a_n\le t\le 1}\frac{\mathbb{G}_n(t)}{t} \to_p 1\,.$$

[Hint: take $\lambda = 1+\epsilon$ with $\epsilon > 0$ in (b) and note that $h(1+\epsilon) > 0$.]

Solution: (a) By the hint, the conditional distribution of $n\mathbb{G}_n(s)$ given $n\mathbb{G}_n(t)$ is Binomial $(n\mathbb{G}_n(t), s/t)$, and thus

$$E\left\{n\mathbb{G}_n(s)\Big|\mathcal{F}_n(t)\right\} = E\left\{n\mathbb{G}_n(s)\Big|n\mathbb{G}_n(t)\right\} = n\mathbb{G}_n(t)\cdot\frac{s}{t} \quad \text{a.s.;}$$

dividing across this identity by $n \cdot s$ yields the claim:

$$E\left\{\frac{\mathbb{G}_n(s)}{s}\Big|\mathcal{F}_n(t)\right\} = \frac{\mathbb{G}_n(t)}{t} \quad \text{a.s..}$$

(b) Since $\{\mathbb{G}_n(t)/t, \mathcal{F}_n(t)\}_{0<t\le 1}$ is a reverse martingale and $\varphi_r(x) \equiv e^{rx}$ is a convex function of x for each $r \in \mathbb{R}$, it follows that $\{\varphi_r(\mathbb{G}_n(t)/t), \mathcal{F}_n(t)\}_{0<t\le 1}$ is a reverse sub-martingale. Thus by Doob's maximal inequality it follows that for $r > 0$ we have

$$\begin{aligned} P\left(\sup_{a\le t\le 1}\frac{\mathbb{G}_n(t)}{t} > \lambda\right) &= P\left(\sup_{a\le t\le 1}\exp\left(r\mathbb{G}_n(t)/t\right) > e^{r\lambda}\right) \\ &\le \frac{E\exp(r\mathbb{G}_n(a)/a)}{\exp(r\lambda)}. \end{aligned}$$

Now the expectation in the numerator on the right side is just

$$\begin{aligned} E\exp\left((r/na)n\mathbb{G}_n(a)\right) &= \left(E\exp((r/na)1_{[0,a]}(\xi_1))\right)^n \\ &= ((1-a)\exp((r/na)\cdot 0) + a\exp((r/na)))^n \\ &= \left((1-a) + ae^{r/(na)}\right)^n \\ &= \left(1 - a(1-e^{r/(na)})\right)^n \\ &\le \exp\left(-na(1-e^{r/(na)})\right) \end{aligned}$$

Combining this with the denominator gives

$$\begin{aligned} P\left(\sup_{a\le t\le 1}\frac{\mathbb{G}_n(t)}{t} > \lambda\right) &\le \exp\left(-na(1-e^{r/(na)}) - r\lambda\right) \\ &= \exp\left(-na(1-e^{r/(na)} + (r/na)\lambda)\right). \end{aligned}$$

for each $r > 0$. Minimizing this with respect to $r > 0$ we find that the minimizing value of r is $r^* \equiv na\log\lambda$. The resulting bound is found to be

$$P\left(\sup_{a\le t\le 1}\frac{\mathbb{G}_n(t)}{t} > \lambda\right) \le \exp(-nah(\lambda))$$

(c) The function h is strictly convex with $h(1) = 0$, $h(0) = 1$, $h'(\lambda) = \log(\lambda)$ and $h''(\lambda) = 1/\lambda > 0$, and therefore $h(1+\epsilon) > 0$ for each $\epsilon > 0$. Thus for each fixed $\epsilon > 0$ we find that

$$P\left(\sup_{a_n\le t\le 1}\frac{\mathbb{G}_n(t)}{t} > 1+\epsilon\right) \le \exp(-na_nh(1+\epsilon)) \to 0$$

if $na_n \to \infty$. Since $\sup_{a_n\le t\le 1}(\mathbb{G}_n(t)/t) \ge \mathbb{G}_n(1)/1 = 1$, it follows that $\sup_{a_n\le t\le 1}(\mathbb{G}_n(t)/t) \to_p 1$.

7. (36 points) Let $p \in (0,1)$. Suppose that $\{Z_n\}_{n=0}^{\infty}$ is a sequence of random variables with

$$P(Z_{n+1} = 2j | Z_n = i) = \binom{i}{j} p^j (1-p)^{i-j}, \quad j \in \{0, \ldots, i\}, i \in \{0, 2, 4, \ldots\}. \tag{2}$$

with the convention that $P(Z_{n+1} = 0 | Z_n = 0) = 1$. Also assume that $P(Z_0 = k_0) = 1$ for a fixed (possibly large) even integer $k_0 \geq 2$. Let $\mathcal{A}_n \equiv \sigma[Z_0, Z_1, \ldots, Z_n]$, $n \geq 1$.
(a) Interpret the sequence $\{Z_n\}$ in terms of Binomial and Bernoulli random variables.
(b) Show that when $p = 1/2$ the process $\{Z_n, \mathcal{A}_n\}$ is a martingale with mean k_0 (with respect to the filtration $\{\mathcal{A}_n\}$ with $\mathcal{A}_n = \sigma\{Z_0, \ldots, Z_n\}$ for $n \geq 0$). For what values of p is $\{Z_n, \mathcal{A}_n\}$ a sub-martingale? For what values of p is it a super-martingale?
For the rest of the problem, assume that $p = 1/2$**.**
(c) Show that with $Y_n \equiv P(Z_{n+1} = 0 | \mathcal{A}_n) = P(Z_{n+1} = 0 | Z_n)$, the process $\{Y_n, \mathcal{A}_n\}_{n\geq 0}$ is a sub-martingale. In fact, use Jensen's inequality to show that $\{Y_n, \mathcal{A}_n\}_{n\geq 0}$ is an almost surely strictly increasing sub-martingale in the sense that $E(Y_{n+1}|\mathcal{A}_n) > Y_n$ almost surely.
Hint: Express Y_n in terms of an exponential function.
(d) Use the result of (d) to show that $Y_n \to_{a.s.} 1$ and hence that $Z_n \to_{a.s.} 0$.

Solution: (a) Note that (2) implies that

$$Z_{n+1} \stackrel{d}{=} 2\text{Binomial}(Z_n, p) \ \text{ conditionally on } \ Z_n.$$

Furthermore, a Binomial(m,p) is equal in distribution to $\sum_{j=1}^m B_j$ where the B_j's are i.i.d. Bernoulli(p) random variables.
(b) It follows easily from (a) that

$$E(Z_{n+1}|\mathcal{A}_n) = E(Z_{n+1}|Z_n) = 2Z_n p \ \text{ a.s.}$$

Thus $\{Z_n, \mathcal{A}_n\}_{n\geq 0}$ is a martingale (with mean $k_0 = Z_0$) if $p = 1/2$, while it is a sub-martingale or super-martingale depending on whether $p > 1/2$ or $p < 1/2$.
(c) From the formula (2) it follows that $Y_n \equiv P(Z_{n+1} = 0|Z_n) = \binom{Z_n}{0} p^0 (1-p)^{Z_n} = \exp(-Z_n \log(1/(1-p)))$ where $\log(1/(1-p)) > 0$ for $p \in (0,1)$. When $p = 1/2$ this becomes $Y_n \equiv \exp(-Z_n \log 2)$. Since the function $\varphi(x) = \exp(-x \log 2)$ is convex and Z_n a non-negative martingale it follows that Y_n is a sub-martingale; in fact, since φ is strictly convex and the Binomial distribution is not concentrated at its mean, $E(Y_{n+1}|\mathcal{A}_n) = E(\exp(-Z_{n+1}\log 2)|\mathcal{A}_n) > \exp(-Z_n \log 2) = (1/2)^{Z_n} = Y_n$ almost surely. Furthermore, note that

$$Es^{\text{Binomial}(n,p)} = (1-p+ps)^n.$$

Thus

$$EY_{n+1} \;=\; E(1/2)^{Z_{n+1}} = E\{E((1/2)^{Z_{n+1}}|Z_n)\} = E\{[(1/2)(1+(1/4))]^{Z_n}\}$$

$$
\begin{aligned}
&= E(s_1^{Z_n}) = E\{[(1/2)(1+s_1)]^{Z_{n-1}}\} \;\text{ with }\; s_1 \equiv (1/2)(1+1/4) \equiv f(1/2) \\
&\qquad \text{where }\; f(s) \equiv (1/2)(1+s^2) \\
&= E s_2^{Z_{n-1}} \;\text{ with }\; s_2 \equiv f(s_1) \\
&= \cdots = s_{n+1}^{Z_0} \quad \text{where }\; s_{n+1} \equiv f(s_n)
\end{aligned}
$$

Thus $s_{n+1} \to s_\infty$ where $s_\infty = f(s_\infty) = (1/2)(1+s_\infty^2)$. Since the only fixed point of f is 1, we find that $s_\infty = 1$ and hence $E(Y_{n+1}) \to 1$.
(d) Thus $\{Y_n, \mathcal{A}_n\}_{n\geq 0}$ is an integrable sub-martingale, and by the s-martingale convergence theorem, $Y_n \to_{a.s.} Y_\infty$. It follows that $Y_\infty = 1$ and this implies that $Z_n \to_{a.s.} 0$.

Remark: Since $Z_{n+1} \stackrel{d}{=} \sum_{j=1}^{Z_n} X_{n,j}$ where the $X_{n,j}$'s are i.i.d. 2Bernoulli(1/2) with $m = E(X_{1,1}) = 1$, this is a special case of the branching process example 13.4.5, page 366, PfS with basic martingale having mean k_0 rather than 1. Note that (c) and (d) give a different proof of Theorem 13.4.1(ii) in this special case.