310 Quals Strategy Compendium

July 13, 2025

1 Conditional Expectation

Definition 1 (Absolute continuous measures)
We say that ν is absolutely continuous wrt μ, ie $\nu \ll \mu$ if

$$\mu(A) = 0 \implies \nu(A) = 0$$

Theorem 2 (Radon-Nikodym Theorem)
If $\nu \ll \mu$, both sigma finite, then there exists $f \in m\mathcal{F}_+$ finite valued such that $\nu = f\mu = \int f d\mu$.

Ie existence of a density wrt dominating measure. f is called the Radon-Nikodym derivative and we write $f = \frac{d\nu}{d\mu}$.

We use the above to prove the existence of CE.

Definition 3 (Conditional Expectation - CE)
Given $X \in L^1(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$, then there exists $Y \in L^1(\Omega, \mathcal{G}, P)$:

$$Y := \mathbf{E}[X|\mathcal{G}] \quad \text{such that} \quad \mathbf{E}[(X - Y)\mathbf{1}_G] = 0 \quad \forall G \in \mathcal{G}.$$

Defined uniquely for P-almost every ω.

To check that something is a CE, use the following:

Theorem 4 (Check CE on π system)
Ex 4.1.3 Dembo. If $\mathcal{G} = \sigma(P)$ and P a π system, then if the above holds for every $G \in P$, then $Y = \mathbf{E}[X|\mathcal{G}]$.
That is, we can just check for every set in a **generating π-system**.

Recipe 5 (Prove something is CE)
Do the following

1. Check that $X \in L^1$.
2. Check that candidate $Y \in L^1$.
3. Show that X and Y integrate $\mathcal{G}$-sets the same, possibly using the pi system fact.

A note of caution that $Y_\pm \neq \mathbf{E}[X_\pm|\mathcal{G}]$ in general.

1.1 Properties of CE

If $X \in L^1(\Omega, \mathcal{F}, P)$:

1. If also $X \in L^1(\Omega, \mathcal{G}, P)$, then $\mathbf{E}X|\mathcal{G} = X$.
2. Drop conditioning if independent:
   (a) if $\mathcal{H} \perp\!\!\!\perp \sigma(\mathcal{X})$, then $\mathbf{E}[X|\mathcal{H}] = \mathbf{E}X$.
   (b) If $\mathcal{H} \perp\!\!\!\perp \sigma(\sigma(X), \mathcal{G})$, then $\mathbf{E}[X|\sigma(\mathcal{H}, \mathcal{G})] = \mathbf{E}[X|\mathcal{G}]$
3. $X \geqslant 0 \implies Y \geqslant 0$ a.s., and $X > 0 \implies Y > 0$ a.s.
4. Linearity
5. Monotonicity: if $X_1 \leqslant X_2$ then $\mathbf{E}X_1|\mathcal{G} \leqslant \mathbf{E}X_2|\mathcal{G}$.
6. If $\mathbf{E}X|Y = Y$ and $\mathbf{E}Y|X = X$ then $X = Y$.
7. Tower
   (a) $\mathbf{E}X = \mathbf{E}[\mathbf{E}(X|\mathcal{G})]$
   (b) If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, then $\mathbf{E}[X|\mathcal{H}] = \mathbf{E}[\mathbf{E}(X|\mathcal{G})|\mathcal{H}]$. "The small one stays on the outside"
8. If $\mathbf{E}[X|\mathcal{G}] \perp\!\!\!\perp X$, then $\mathbf{E}X|\mathcal{G} = \mathbf{E}X$.
9. Take-out: if $Y \in m\mathcal{G}$, then $\mathbf{E}[XY|\mathcal{G}] = Y\mathbf{E}[X|\mathcal{G}]$.
10. Conditional Jensen
11. Conditional Markov, Holder
12. MCT, Fatou, DCT for CE
13. If $X_n \xrightarrow{L^q} X_\infty$, then $\mathbf{E}[X_n|\mathcal{G}] \xrightarrow{L^q} \mathbf{E}[X_\infty|\mathcal{G}]$ (apply Jensen)
14. $\{\mathbf{E}[X|\mathcal{H}] : \mathcal{H} \subset \mathcal{F} \text{ is a sigma algebra}\}$ is UI. (4.2.33).

1.2 CE as Orthogonal Projection

If $X \in L^2$, then $Y = \mathbf{E}[X|\mathcal{G}]$ is the unique $Y \in L^2$ such that $\|X - Y\|_2 = \inf\{\|X - W\|_2 : W \in L^2(\Omega, \mathcal{G}, P)\}$. Ie, the conditional expectation is a projection onto the subspace with respect to the $\langle X, Y \rangle = \int XY dP$ inner product. (Since $L^2(\Omega, \mathcal{G}, P)$ is a *Hilbert Subspace*).

There exists a unique projection in Hilbert Spaces onto subspaces, ie $\langle h - \hat{h}, f \rangle = 0$ where $\hat{h} = \text{Proj}_{L^2(\Omega, \mathcal{G}, P)}\, h$ and $f \in L^2(\Omega, \mathcal{F}, P)$.

Theorem 6 (Cauchy-Schwarz in L^2)

$$|\mathbf{E}XY| \leqslant \sqrt{\mathbf{E}X^2\mathbf{E}Y^2}$$

1.3 Regular conditional probabbility distributions (RCPD)

Theorem 7 (Can take conditional expectation wrt conditional density)
If X, Z have a joint density and $g(X)$ is integrable, then

$$\mathbf{E}[g(X)|Z] = \int_{\mathbb{R}} g(x) f_{X|Z}(x|z) dx$$

as in elementary probability.

Definition 8 (RCPD (Regular conditional probability distribution))
Let Y be $\mathbb{S}$-valued random variable then

$$\hat{P}_{Y|\mathcal{G}}(\cdot, \cdot) : \mathcal{S} \times \Omega \mapsto [0, 1]$$

is the RCPD of Y given $\mathcal{G}$ if:

1. $\hat{P}_{Y|\mathcal{G}}(A, \cdot)$ is a version of the CE $\mathbf{E}[\mathbf{1}[Y \in A|\mathcal{G}]$ for $A \in \mathcal{S}$.
2. For any fixed ω, $\hat{P}_{Y|\mathcal{G}}(\cdot, \omega)$ is a probability measure on $(\mathbb{S}, \mathcal{S})$.

Analogue to Markov Kernel in 310a.

Theorem 9 (RCPD existence)
If X real and $\mathcal{G}$ a $\sigma-$algebra, then RCPD exists.

Also true for any $\mathcal{B}$-isomorphic rv X.

Also used to show existence of transition probability- see Ex 4.4.5.

A helpful exercises 4.4.6 shows we can calculate expectations using the RCPD:

$$\mathbf{E}[h(X, Y)|\mathcal{G}](\omega) = \int_{\mathbb{R}} h(x, Y(\omega)) d\hat{P}_{X|\mathcal{G}}(x, \omega)$$

ie fix ω and integrate over $x \in \mathbb{R}$.

2 Martingales

To check a martingale:

1. X_n is integrable for all n
2. Adapted
3. $\mathbf{E}[X_{n+1}|\mathcal{F}_n] = X_n$ for all n

Alternatively, if $X_n = \sum_{k=0}^{n} D_k$, then check that $\mathbf{E}[D_{n+1}|\mathcal{F}_n] = 0$.

Example 10 (Quadratic martingale)
If $\mathbf{E}\xi_i = 0$ and $\mathbf{Var}\xi_i = \sigma^2 < \infty$ and the ξ_i are independent, then

$$S_n^2 - n\sigma^2 \quad \text{is a martingale}$$

Example 11 (Exponential martingale)
If S_n random walk with independent, iid increments,

$$M_n = \prod_{i=1}^{n} \exp(\theta \xi_i)/\phi(\theta) = \exp(\theta S_n)/\phi(\theta)^n$$

Is a special case of product martingale

A predictable sequence of random variables A_n gives the amount of money you'd be willing to bet at time n– must be based on information from previous time points, up through $n-1$, to make the n-th bet. Think of A_n as your n-th bet- given by information from time $\{1, \ldots, n-1\}$. Can think of winnings in the following decomposition:

$$\sum_{m=1}^{n} H_m(X_m - X_{m-1})$$

where H_m is the amount you wager between days m and m+1 X_m is the stock price. So our profit is the difference in prices times the amount that we bet/number of shares we hold.

The above is known as the martingale transform of $\{X_n\}$ by the predictable process $\{A_n\}$.

Fact 12 (Martingale transforms)
MG transforms of mgs $\implies$ mg

Mg transforms of sub/sup mgs $\implies$ sub/sup mg

Example 13 (Conditions of a.s. convergence of (sub/sup) mg don't necessarily give L^1 convergence)
Durrett 193 - example with S_n simple random walk *starting at* $S_0 = 1.$, let $X_n = S_{\tau \wedge n}$ where τ is first time $S_n = 0$. Then can X_n is a non-negative martingale so a.s. limit exists, must be $X_n \xrightarrow{\text{a.s.}} 0$. But $\mathbf{E}X_n = 1$. So L^1 convergence can't occur.

Example 14 (Martingale converging as to $-\infty$)
Construct something such that the positive event happens only finitely often.. $X_n \xrightarrow{\text{a.s.}} -\infty$ even though X_n is a fair bet. There's a remote chance of a big reward.

Theorem 15 (Martingale with bounded increments converges or oscillates between $\pm\infty$)
(Durrett 4.3.1)
If X_n mg with $|X_{n+1} - X_n| \leqslant M < \infty$, then if

$$C = \{\lim X_n < \infty \text{ exists}\} \quad D = \{\limsup X_n = \infty, \liminf X_n = -\infty\}$$

Then:

$$P(C \cup D) = 1.$$

So given a martingale, any ω in a set of prob 1 must be do one of these two things.

Example 16 (Biased random walk)
If positive step is $p \neq 1/2$, then

$$X_n = \left(\frac{q}{p}\right)^{S_n} \quad \textbf{is a martingale}$$

Theorem 17 (Wald's Equation)

Theorem 18 (Doob Decomposition)
For any $\{X_n\}$ stochastic process adapted to $\{\mathcal{F}_n\}$, write:

$$M_n = \sum_{k=0}^{n-1}(X_{k+1} - \mathbf{E}(X_{k+1}|\mathcal{F}_k)) \quad A_n = \sum_{k=0}^{n-1}(\mathbf{E}(X_{k+1}|\mathcal{F}_k) - X_k)$$

so that

$$X_n = X_0 + M_n + A_n,$$

where M_n is a martingale and $A_n \in m\mathcal{F}_{n-1}$ ie is adapted. In summary, an adapted (discrete) stochastic process can be written as the sum of a martingale and a predictable process.

In the case that X_n is a submartingale, A_n is non-negative and increasing.

Theorem 19 (BC2 version 2)
If $\{B_n\}$ sequence of events, then:

$$\{B_n \text{ i.o.}\} = \{\sum_{n=1}^{\infty} P(B_n|\mathcal{F}_{n-1}) = \infty\}.$$

Proof. Idea is that $X_n = \sum_{m \leqslant n} \mathbf{1}[B_m]$ is a submartingale, apply the Doob's decomposition and then note that:

$$|M_n - M_{n-1}| \leqslant 1$$

so we can apply the theorem with bounded martingale differences to get that $P(C \cup D) = 1$, show that this means that the two events are the same. □

Note that Dembo has an extra comment about the rate at which X_n goes to infinity.

Example 20 (Polya's Urn)
(Durret section 4.3.2)
Contains r red and g green balls– each time we draw a ball, we replace it and add c more of the same color. Let G_n is # of green balls, X_n is the *fraction* of green balls after the n-th draw, ie G_n/N_n.

1. X_n is a non-negative martingale
2. $X_n \xrightarrow{\text{a.s.}} X_\infty$ as
3. If $b = g = 1$, then $P(G_n = m+1) = \frac{1}{n+1}$, ie uniform on $\{1, \ldots, m+1\}$.
4. X_∞ then has a uniform distribution on $(0, 1)$.
5. In general, $X_\infty \stackrel{d}{=} \text{Beta}(g/c, r/c)$.

Example 21 (Likelihood ratios)
Durrett..

2.1 Branching process

If $\mu = \mathbf{E}\xi_i^{(m)}$ where the $\xi_i^{(m)}$ are the number of offspring of the i-th member of the m-th generation, then:

$$\boxed{Z_n/\mu^n \text{ is a martingale}}$$

If $\mu < 1$ or $\mu \leqslant 1$ and $P(\xi_i = 1) < 1$, then the population dies out, ie $Z_n/\mu^n \xrightarrow{\text{a.s.}} 0$. "If the average number of offspring is fewer than 1, then the population dies out. For $\mu > 1$, we use *generating functions* and we can prove things about the limiting random variables.

2.2 Doob's maximal inequality and L^p convergence

Think of Doob's inequality as an improvement on Markov's for (sub)-martingales.

Theorem 22 (Doob's Inequality)
For submartingale $\{X_n\}$, we have:

$$P(\max_{k \leqslant n} X_k \geqslant \lambda) \leqslant \lambda^{-1} \mathbf{E} X_n^+.$$

Kolmogorov's Maximal inequality is a special case. If $Z_n = \sum_{i=1}^n Y_i$ with $\mathbf{E}Y_i = 0$, $\mathbf{Var}Y_i < \infty$, then Z_n is obviously a mg and

$$P(\max_{k \leqslant n} |Z_n| > \lambda) = P(\max_{k \leqslant n} Z_n^2 > \lambda^2) \leqslant \frac{1}{\lambda^2} \mathbf{Var} Z_n.$$

Theorem 23 (L^p maximal inequality for $p > 1$)
If $\{X_n\}$ submg,

$$\mathbf{E}(\max_{k \leqslant n}(X_n^p)_+) \leqslant \left(\frac{p}{p-1}\right)^p \mathbf{E}(X_n^+)^p.$$

So if $\{X_n\}$ is a mg,

$$\mathbf{E}(\max_{k \leqslant n}|X_n|)^p \leqslant \left(\frac{p}{p-1}\right)^p \mathbf{E}|X_n|^p.$$

The previous L^p inequality leads to:

Theorem 24 (L^p convergence theorem)
If X_n is a *martingale* with $\sup \mathbf{E}|X_n|^p < \infty$ with $p > 1$, then $X_n \xrightarrow{\text{a.s.}} X$ and $X_n \xrightarrow{\text{L}^\text{p}} X$.

Proof. Almost sure convergence comes from the Doob's convergence theorem and Markov. L^p convergence comes from L^p maximal inequality, see Durrett 4.4.6. □

Note that this theorem:

1. Is only for martingales
2. Does not have a L^1 analog

2.3 Square Integrable Martingales

Fact 25 (Square Integrable martingales (Dembo Ex 5.1.8))
Have uncorrelated differences D_n. Durrett 4.4.7.

Suppose that X_n is an L^2 martingale with $X_0 = 0$. Then:

1. X_n^2 is a submartingale
2. Apply Doob's Decomposition theorem $X_n^2 = M_n + A_n$ with M_n mg and A_n is a **predictable, increasing** sequence, called the **predictable compensator**.
3. $\langle X \rangle_n := A_n = X_0^2 + \sum_{m=1}^n \mathbf{E}(X_m^2|\mathcal{F}_{m-1}) - X_{m-1}^2 = \sum_{m=1}^n \mathbf{E}((X_m - X_{m-1})^2|\mathcal{F}_{m-1})$
4. Think of the predictable compensator as the total variance at time n of the path made by $\{X_n\}$.

Theorem 26 (Finite predictable compensator limit means finite limit of the original martingale)
If $\langle X \rangle_\infty < \infty$ then we have that $X_n \xrightarrow{\text{a.s.}} X_\infty < \infty$.

Recipe 27 (Convergence of random series)
Dembo 5.3.37. Want to show that $\sum_{n=1}^\infty X_n(\omega)$ converges. Do we need symmetric distribution of X_n

Example 28 (Quals second q this week)
...