MEASURE THEORETICAL PROBABILITY I HOMEWORK 7

TOMMENIX YU
STAT 38100
DUE MAR 3RD

Discussed with classmates.

Exercise 1. *Prob 1.*

Proof.

Show that it is a polished space:

We first show that it is indeed a metric, then the space is complete, then the space is separable.

It's a metric: Note that it's well defined since the sum is less than the geometric series $\sum_{i=0}^{\infty} 2^{-n}$, as each coefficient is less than 1.

- (Positive definite): Since each term in the summand is non-negative, the sum is non-negative. Moreover, when the sum is 0 we know that each term is 0, which means that f and g agrees on each length 1 interval, hence equal.
- (Symmetric): since $|f-g|_{j,j+1} = |g-f|_{j,j+1}$ we know that the metric is symmetric.
- (Triangle inequality) We show the triangle inequality on each length 1 interval, then it holds if we sum up all such intervals, provided the summation converges, which it does. Now denote $a := |f-g|_{j,j+1}, b := |f-h|_{j,j+1}, c := |g-h|_{j,j+1}$ then we have by the fact that sup norm is a norm on any interval:

$$a + c - b \geq 0$$

$$\Rightarrow a + c - b + 2ac + abc \geq 0$$

$$\Rightarrow a + c + 2ac + ab + cb + 2abc \geq b(1+a)(1+c)$$

$$\Rightarrow \frac{a}{1+a} + \frac{c}{1+c} \geq \frac{b}{1+b}$$

which is what we want.

Completeness:

To show completeness we show that for any Cauchy sequence in the space, the sequence converges.

Now, if a sequence f_n is Cauchy, then for any $\varepsilon > 0$ there exitst N such that for all $n, m > N$ we know

$$\sum_{j=0}^{\infty} 2^{-j} \frac{|f_n - f_m|_{j,j+1}}{1 + |f_n - f_m|_{j,j+1}} < \varepsilon$$

which in particular means $\forall j \in \mathbb{N}^*$ such that $2^j \varepsilon < 1$:

$$2^{-j} \frac{|f_n - f_m|_{j,j+1}}{1 + |f_n - f_m|_{j,j+1}} < \varepsilon$$

$$\Rightarrow |f_n - f_m|_{j,j+1} \leq \frac{2^j \varepsilon}{1 - 2^j \varepsilon}.$$

This means that for all $[0, j]$, the difference between f_n and f_m is less than $\dfrac{2^j \varepsilon}{1 - 2^j \varepsilon}$.

Now we find a candidate for the limit, then show that it actually is the limit.

For each $x \in [0, \infty)$ we know that $f_n(x)$ is a Cauchy sequence due to above relation, thus it converges in $\mathbb{R}$. Define f to be the point wise limit of f_n in the above way. We have to check that $f \in C[0, \infty)$. But we know that for any M, $f_n \to f$ uniformly on $[0, M]$, plus f_n are continuous, and hence f is continuous on $[0, M]$. So f is not discontinuous at any point since every point in the real line lies within such an interval, so $f \in C[0, \infty)$.

Now we show that f is the norm limit of f_n. For any small ε (for large we use any smaller ε), we find

$$N = N_\varepsilon := \sup \left\{ 2^N \varepsilon < 1 \middle| N \in \mathbb{N} \right\}$$

then we have that for any $n \geq N$

$$\begin{aligned} \rho(f_n, f) &= \sum_{j=0}^{N-1} 2^{-j} \frac{|f_n - f|_{j,j+1}}{1 + |f_n - f|_{j,j+1}} + \sum_{j=N}^{\infty} 2^{-j} \frac{|f_n - f|_{j,j+1}}{1 + |f_n - f|_{j,j+1}} \\ &\leq N\varepsilon + 2^{N-1} \end{aligned}$$

where the first bound is by above deduction for each j, and for the second we bound the fraction by 1. Yet since $2^N \varepsilon < 1$ we have

$$2^N \varepsilon \frac{N}{N} \leq 1 \Rightarrow N\varepsilon \frac{2^N}{N} < 1$$

and since $N \to \infty$ as $\varepsilon \to 0$, we have $\dfrac{2^N}{N} \to \infty$ and thus $N\varepsilon \to 0$ as $\varepsilon \to 0$. This gives us that

$$0 \leq \rho(f_n, f) \leq N\varepsilon + 2^{N-1} \to 0$$

which means convergence in norm.

Separable:

By Stone Weierstrass we know that any continuous function on an interval can be approximated by the class $PQ[a, b]$ of polynomials with rational coefficients, which is thus countable. Now we define

$$D := \bigcup_{n=1}^{\infty} PQ[0, n]$$

then it's the class of polynomials with rational coefficients supported on all intervals of the type $[0, n]$ for $n \in \mathbb{N}^+$. It's countable since countable union of countable sets is countable.

Now, for any function $f \in C[0, \infty)$, we construct sequences of functions in $PQ[0, n]$ to approximate $f\Big|_{[0,n]}$ for each n. Call these sequences $\{g_{1,n}\}, \{g_{2,n}\}, \dots$. Then we pick the diagonal sequence

$$g_n = \{g_{n,n}\}$$

then we know that for any $\varepsilon > 0$ we pick the N_ε in the exactly same manner as we've chosen above, and get

$$\begin{aligned}\rho(g_n, f) &= \sum_{j=0}^{N-1} 2^{-j} \frac{|g_n - f|_{j,j+1}}{1 + |g_n - f|_{j,j+1}} + \sum_{j=N}^{\infty} 2^{-j} \frac{|g_n - f|_{j,j+1}}{1 + |g_n - f|_{j,j+1}} \\ &\leq N\varepsilon + 2^{N-1} \to 0\end{aligned}$$

for all large enough $n > M$ since we've uniformly approximated f on $[0, M] \supset [0, N]$. Thus, since $g_n \in D$ we have that D is a countable dense set of $C[0, \infty)$, so the space is separable.

In conclusion, this is a polish space.

<u>$f_n \to f \iff f_n \to f$ uniformly on compact sets:</u>

We note that for the backwards direction, we use the compact sets $[0, M]$ and an exact same argument (separable part, note that is exactly what we assumed) as above to show convergence in norm ρ. So we are only left with the forward direction.

As for this direction, for any compact subset of $\mathbb{R}$ we know that there is an M such that $M > |x|$ for all x in the compact set (otherwise no finite cover). So we only need to prove for all $[0, M]$, $f_n \to f$ uniformly. Now since $f_n \to f$ we know

$$\sum_{j=0}^{\infty} 2^{-j} \frac{|f_n - f|_{j,j+1}}{1 + |f_n - f|_{j,j+1}} < \varepsilon$$

for all ε as $n \to \infty$. This in particular means $\forall \delta > 0, \forall j \in \mathbb{N}^*$, we can choose ε small such that $2^j \varepsilon < 1$ and $\dfrac{2^j \varepsilon}{1 - 2^j \varepsilon} < \delta$ we have

$$2^{-j} \frac{|f_n - f_m|_{j,j+1}}{1 + |f_n - f_m|_{j,j+1}} < \varepsilon$$

$$\Rightarrow |f_n - f_m|_{j,j+1} \leq \frac{2^j \varepsilon}{1 - 2^j \varepsilon} < \delta$$

where this means

$$|f_n - f_m|_{0,j+1} < \delta$$

which gives uniform convergence on all $[0, M]$. By arguments above, this gives the uniform convergence on all compact sets.

Every open ball is in the π σ-algebra:

The main approximation of this proof is (details see below steps):

(1) We use close set from below to approximate open balls, this gets us to approximation with evaluation at rational points.
(2) We use open sets within the above approximation to get down to a closed set with intersections, this gets us to the closure of all choice of sequences b_j using rational sequences c_j^k that helps us get the norm convergence.

Step 1: reduce to closed balls:

Fix ε. To show this, we use the dense set D of $C[0, \infty)$ defined above. So for any $B_f(\varepsilon)$ we know that $D \cap B_f(\varepsilon)$ is dense in $B_f(\varepsilon)$. More over, we note that

$$B_f(\varepsilon) = \bigcup_{n=N}^{\infty} B_f\left(\varepsilon - \frac{1}{n}\right)$$

for large enough N such that $\varepsilon - \dfrac{1}{N} > 0$. Thus, let $\varepsilon' \in \left\{\varepsilon - \dfrac{1}{n}\middle| n \geq N\right\}$, we only need to show that the closed ball $\overline{B_f(\varepsilon')}$ is in the σ-algebra.

The reason for this toil is because we note that even if $|f - g| < \varepsilon$ for all rational points in an interval, and f, g continuous, we still can only get $|f - g| \leq \varepsilon$ on the whole interval. So we reduce the problem to closed balls, which is easier to maneuver. Moreover, this helps us to take the closure of any dense subset.

Step 2: reduce to only rational points, over rational combinations:

The true claim here is that

$$S := \overline{\left\{ g \middle| |g(q_j) - f(q_j)| \leq c_j \in \mathbb{Q}^+; \forall \sum_{j=0}^{\infty} \frac{c_j}{2^j} < \varepsilon'; \forall q_j \in \mathbb{Q} \cap [j, j+1]; f, g \in C[0, \infty) \right\}}$$
$$= \overline{B_f(\varepsilon')}$$

What this S is, is that for all positive rational sequence c_j satisfying $\displaystyle\sum_{j=0}^{\infty} \frac{c_j}{2^j} < \varepsilon'$, if all of the rational numbers in the corresponding $[j, j+1]$ has that $|g(q_j) - f(q_j)| \leq c_j$, then we count that g as a member of S.

$\subset$:

Note that we only need to show the non-closure version of S is contained.

Since g is continuous, we assume that $g \notin \overline{B_f(\varepsilon')}$, then there is a point g such that $|g(x) - f(x)| > c_j$ where $x \in [j, j+1]$ (note here if $x = j$ then we're automatically done since $x \in \mathbb{Q}$ and by definition), but this is impossible since g is sequentially continuous and we have a $\leq c_j$ sequence of rational numbers that converges to x. Contradiction! So this direction holds.

$\supset$:

For this direction the main part is to pass from all real b_j with $\sum_{j=0}^{\infty} \frac{b_j}{2^j} < \varepsilon'$ to rational c_j such that $\sum_{j=0}^{\infty} \frac{c_j}{2^j} < \varepsilon'$. This again can be done with a rational approximation of each b_j by c_j^k from below, as all such c_j^k satisfies the convergent condition. This approximation plus the closure yields the result.

The rest are by definition.

Step 3: S with out closure is in the σ-algebra:

Note that the cross product of closed boxes $\prod_{i=1}^{n} [f(t_1 - c_j), f(t_1 + c_j)] \in \mathcal{B}(\mathbb{R})$ where we denote

$$T := \{t_1, \dots, t_n\} \in \left\{ \{t_1, \dots, t_n\} \middle| n \in \mathbb{N}^+ \right\} =: \mathcal{T}.$$

Now, we note that there's only countable rational numbers, so we can just let $n = 1$ and use the intersection

$$S' := \bigcup_{c_j} \bigcap_{i=1}^{\infty} \pi_{\{q_i\}}^{-1} \left([f(q_i - c_j), f(q_i + c_j)] \right)$$

where $q_1, q_2, \dots$ is a reordering of $\mathbb{Q}^+$. Note that this holds for all validate rational sequences c_j.

We note first that $S' \in \sigma \left\{ \pi_T \middle| T \in \mathcal{T} \right\}$ by definition.

Then, we also note that S' is indeed the part of S without closure.

Step 4: pass the closure:

We go back to the content of step 1 again. We want to be able to get around the closure, i.e. to pass to all such sequences b_j from c_j. This requires us to note that we can use $\delta = \frac{1}{2}(\varepsilon + \varepsilon')$ as our representative for $B_f(\delta)$. What this extra $\frac{1}{2}(\varepsilon - \varepsilon')$ does is that it can be distributed to each $[j, j + i]$ with weight 2^{-j} as a buffer for existence of a decreasing rational sequence c_j that limits to b_j. Once this is done, we reform our S' above with δ and add an additional layer of intersection to close the set, that is, we define

$$S'' := \bigcap_{\mathcal{A}_m} \bigcup_{c_j(\delta)} \bigcap_{i=1}^{\infty} \pi^{-1}_{\{q_i\}} \left([f(q_i - c_j), f(q_i + c_j)] \right)$$

where the index set $\mathcal{A}_m := \delta_m - \epsilon' \leq \frac{1}{m}$ where $\delta_m \in [\epsilon', \delta]$ starting with large m.

Now we can say that $S'' = S$ since it is the infinite intersection of open sets from outside.

Step 5: conclusion

Since S'' is in the σ-algebra generated by all finite dimensional projections, and $S = S''$, so is S, which is the closed balls defined above. But by step 1 this is equivalent to show for any open balls, thus we are done.

Now, since all open balls are generated, we know that the Borel algebra is contained. But note that we generate things from Borel sets in $\mathbb{R}^k$ so we really cannot generate any extra sets. i.e. if any non-Borel set is contained then it corresponds to a non-Borel set $\in \mathbb{R}$ in the image of π_x, which is a contradiction. So it generates the Borel σ-algebra.

$\square$

Exercise 2. *Prob 2.*

Proof.

What we want to show is that for any $A \in \mathcal{B}(S)$, $f^{-1}(A)$ is measurable. But it suffices us to show for all $A \in B(S)$, where $B(S)$ is a subbasis of S, since the topology generated by $B(S)$ is $\mathcal{T}(S)$ and hence the σ-algebra generated is $\mathcal{B}(S)$.

But because a polish space is also metrizable, we use the radius balls around all points in S as the subbase of topology: we use $B_x(r)$ for all $r \geq 0$ as the basis.

Now, for any $A := B_x(r) \subset \mathcal{B}(S)$, we know that

$$f^{-1}(A) \subset \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n} f_m^{-1}(A)$$

where the subset relation is often a proper subset relation, since points that lies with in A, an open set, can converge to the boundary of A. But we turn this into a decent equality with the same technique used above i.e. using $A_k := B_x\left(r - \frac{1}{k}\right)$ to exclude all those. That is, we have

$$f^{-1}(A) = \bigcup_{k>N} \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n} f_m^{-1}(A_k)$$

where N is chosen such that $r > \dfrac{1}{N}$. Now we conclude that $f^{-1}(A)$ is measurable, but since $\sigma(A) = \sigma(\mathcal{T}(S))$ we get that f is measurable. $\square$

Exercise 3. *Prob 3.*

Proof.

d is a metric:

Positive definiteness:

By definition, the metric is non-negative since for it's defined as infimum of positive numbers, and 1 is always an upper bound, since the measures are probability measures.

To show $d(\mu, \nu) = 0 \iff \mu = \nu$, we first assume $\mu = \nu$, then we get that for all $A \in \mathcal{B}(S)$, $\mu(A) = \nu(A)$ and hence $\mu(A) \leq \nu(A^0) + 0$ and $\nu(A) \leq \mu(A^0) + 0$, so we are done with the backwards direction.

To show the forward direction, assume $d(\mu, \nu) = 0$. That is, there exists a sequence of real number $a_n \to 0$ such that

$$\mu(A) \leq \nu(A^{a_n}) + a_n \text{ and } \nu(A) \leq \mu(A^{a_n}) + a_n$$

for all $A \in \mathcal{B}(S)$.

Note that for all A closed, we have

$$\mu(A) \leq \nu(A^{a_n}) + a_n \Rightarrow \mu(A) \leq \liminf_{n\to\infty} \nu(A^{a_n}) + a_n = \nu(\overline{A}) = \nu(A)$$

and a reverse argument holds similarity, so we have

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

that is, μ and ν agrees on closed sets. But this means that they agrees on all open sets (complement of closed sets), hence on all $\mathcal{B}(S)$, so $\mu = \nu$.

Symmetric:

Note that the definition really is irrelevant to the order of the measure's appearance, so it's symmetric.

Triangle inequality:

Note that if for fixed ε, δ satisfies that for all $A \in \mathcal{B}(S)$

$$\begin{cases} \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \\ \nu(A) \leq \eta(A^\delta) + \delta \text{ and } \nu(A) \leq \eta(A^\delta) + \delta \end{cases}$$

and thus

$$\mu(A) \leq \eta(A^{\varepsilon+\delta}) + \varepsilon + \delta \text{ and } \eta(A) \leq \mu(A^{\varepsilon+\delta}) + \varepsilon + \delta$$

where we used $A^{\varepsilon+\delta} \supset A^\varepsilon$. But this immediately means the least value that satisfies the two inequalities for μ and η is smaller than the sum of those for μ and ν, and ν and η, since we can find two sequences for each and get that less or equal to is also passed.

Thus, the triangle inequality holds.

$\mu_n \rightharpoonup \mu \iff d(\mu_n, \mu) \to 0$:

$\Rightarrow$:

What we want to show is that $d(\mu_n, \mu) \to 0$. But for the exact same reason as in the positive definite part above, we only need to show for all A closed subset of S, we have

$$\mu(A) \leq \lim_{n\to\infty} \mu_n(\overline{A}) \textit{ and } \lim_{n\to\infty} \mu_n(A) \leq \mu(\overline{A})$$

where since $A = \overline{A}$ we only need

$$\mu(A) \leq \lim_{n\to\infty} \mu_n(A) \textit{ and } \lim_{n\to\infty} \mu_n(A) \leq \mu(A).$$

Now by the Portmanteau Lemma in class where we use the 6th condition that $\mu_n \rightharpoonup \mu \iff$ for every Borel $B \subset S$ such that $\mu(\partial B) = 0$, we have

$$\lim_{n\to\infty} \mu_n(B) = \mu(B).$$

But since here A is closed so $\partial A = \emptyset$ so it has measure 0, thus $\mu(A) = \lim_{n\to\infty} \mu_n(A)$ and hence we are done.

$\Leftarrow$:

If we assume that $d(\mu_n, \mu) \to 0$, we have that for all closed set A that

$$\mu(A) \leq \lim_{n\to\infty} \mu_n(\overline{A}) = \lim_{n\to\infty} \mu_n(A) \textit{ and } \lim_{n\to\infty} \mu_n(A) \leq \mu(\overline{A}) = \mu(A)$$

for which we in particular has

$$\limsup_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \mu_n(A) \leq \mu(A)$$

and by the 4th equivalent condition in the Portmanteau Lemma we're done.

(As reference) The 4th condition is: $\mu_n \rightharpoonup \mu \iff$ for every closed set $F \subset S$, we have

$$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F).$$

$\square$

Exercise 4. *Prob 4*

Proof.

We adopt the definition in Durret's textbook, that a Brownian motion is a real-valued process such that

(1) For $t_0 < t_1 < \cdots < t_n$, then $B(t_0), B(t_1) - B(t_0), \ldots, B(t_n) - B(t_{n-1})$ are independent.
(2) If $s, t \geq 0$, then
$$P(B(s+t) - B(s) \in A) = \int_A (2\pi t)^{-1/2} \exp(-x^2/2t) dx.$$
(3) With probability 1, $t \to B_t$ is continuous.

$-B$ is also a Brownian motion:

We check:

(1) For $t_0 < t_1 < \cdots < t_n$, then $-B(t_0), -B(t_1) + B(t_0), \ldots, -B(t_n) + B(t_{n-1})$ are independent, because $-B(s)$ and $B(s)$ has the same generated σ-algebra.
(2) If $s, t \geq 0$, then
$$P(-B(s+t) + B(s) \in A) = \int_A (2\pi t)^{-1/2} \exp(-x^2/2t) d(-x) = \int_A (2\pi t)^{-1/2} \exp(-x^2/2t) dx$$
since the integrand is even in x.
(3) With probability 1, $t \to -B_t$ is continuous, since it is the combination of (-1) and B_t, both continuous.

Fix $s \geq 0$, $W(t) = B(s+t) - B(s)$ is also a Brownian motion:

We check:

(1) For $t_0 < t_1 < \cdots < t_n$, then $W(t_0) = B(s+t_0) - B(s), W(t_1) - W(t_0) = B(s+t_1) - B(s+t_0), \ldots, W(t_n) - W(t_{n-1}) = B(s+t_n) - B(s+t_{n-1})$ are independent, because $B(s)$ is a constant, and the rest are independent if we pick $s_0 = s + t_0, \ldots$ and use the independent property for $B(s_i)$.
(2) If $u, t \geq 0$, then
$$P(W(u+t) - W(u) \in A) = P(B(u+t+s) - B(u+s) \in A) = \int_A (2\pi t)^{-1/2} \exp(-x^2/2t) dx$$
since the difference of the 2 points B evaluates has difference t.
(3) With probability 1, $t \to B(s+t) - B(s)$ is continuous since $-B(s)$ is a constant and $B(s+t)$ is constant in t.

Fix $c > 0$, $W(t) = c^{-1/2} B(ct)$ is also a Brownian motion:

We check:

(1) For $t_0 < t_1 < \cdots < t_n$, then $W(t_0) = c^{-1/2}B(ct_0), W(t_1) - W(t_0) = c^{-1/2}B(ct_1) - c^{-1/2}B(ct_0), \dots, W(t_n) - W(t_{n-1}) = c^{-1/2}B(ct_n) - c^{-1/2}B(ct_{n-1})$ are independent, because $B(s)$ is a constant since first the multiplication doesn't matter since multiplication is a 1-1 onto map from $\mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R})$, and the same also applies to the $c^{-1/2}$ outside.

(2) If $s, t \geq 0$, then

$$P(W(s+t) - W(s) \in A) = P(c^{-1/2}B(ct+cs) - c^{-1/2}B(cs) \in A)$$

$$= \int_{c^{1/2}A} (2\pi ct)^{-1/2} \exp(-x^2/2ct)dx \overset{y=c^{-1/2}x}{=} \int_A (2\pi t)^{-1/2} \exp(-y^2/2t)dy$$

is the same expression.

(3) With probability 1, $t \to c^{-1/2}B(ct)$ is continuous since $c^{-1/2}$ is a constant and $B(ct)$ is the stretching of a continuous function in t, hence still continuous.

$\square$

Exercise 5. *Prob 5*

Proof.

The limit of finite dimensional distributions converge to that of a Brownian motion:

If the point t is a point that is a multiple of a degree of 2, that is, $t = c \cdot 2^{-s}$ for odd c, then we know that the pointwise sequence $X_n(t)$ is a constant after s terms by definition, so it's corresponding probability is

$$\mathbb{P}(X_n(t) \in A) = \mathbb{P}\left(\sum_{i=1}^{s} a_i N(0, 2^{-i-1}) \in A\right)$$

where a_i is some weight associated to the point's actual position in the interval. Hence, since it's the sum of independent Gaussians, it's still Gaussian.

Now we show that the coefficient a_i is really related to the standard Gaussian. That is, we notice that the a_i are $\frac{d}{2^i}$ where d is the difference between t and its closest $\frac{l}{2^i}$ for l odd. But this coincides exactly with the fact that the variance is additive ($\mathrm{Var}(X+Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y$) and hence we know that

$$\mathbb{P}\left(\sum_{i=1}^{s} a_i N(0, 2^{-i-1}) \in A\right) = \mathbb{P}\left(N(0,t) \in A\right)$$

which means we've checked difference of $X(s)$ and $X(0)$. But this is enough to generate all $X(t)$ and $X(s)$ by additivity of Gaussians.

This concludes for all points that can be finitely expressed. But for any point that is not a grid point (of degrees of 2s), we can list a sequence of odd multiples of 2 that is closest to it, then since the sum converges absolutely we can write out

$$\mathbb{P}\left(\lim_{n\to\infty} X_n(t) \in A\right) = \mathbb{P}\left(\sum_{i=1}^{\infty} a_i N(0, 2^{-i-1}) \in A\right) = \mathbb{P}\left(N(0,t) \in A\right)$$

where the infinite sum is a sum of independent Gaussians that are well defined and converges, thus it's still a Gaussian. Moreover, the same decomposition is contained as in the finite end case since the sum converges.

But then the finite dimensional distribution is

$$\mathbb{P}\left(\lim_{n\to\infty} X_n(t_1) \in A_1, \dots, \lim_{n\to\infty} X_n(t_m) \in A_m\right) = \int_{A_1\times\cdots\times A_n} dN$$

where

$$N = N(0,t_1) \times \cdots \times N(0,t_m)$$

is the multivariate Gaussian with the correct variance, so it's the same as that of a Brownian motion.

Expectation $< \infty$:

The expectation of the sup norm is the maximal difference of differences at each odd multiple of 2^{-k}. In other words, if there are l odd multiples at step n, then the difference is

$$\mathbb{E}\left[||X_{n+1} - X_n||_\infty\right] = \mathbb{E}\left[\max\left\{|Y_1|, \dots, |Y_l|\right\}\right]$$

where Y_i are iid with $Y_1 \sim N(0, 2^{-n})$, so we just have to compute the order distribution of normal distributions and show that they indeed forms a convergent sequence.

To do this, let $Y = \max\{X_1, \dots, X_n\}$, (where just for convenience we use X to stand for Gaussians) then we have by Jensen's inequality that

$$\exp(t\mathbb{E}[Y]) \leq \mathbb{E}[e^{tY}] = \mathbb{E}[\max e^{tX_i}] \leq \sum_{i=1}^{n} \mathbb{E}[e^{tX_i}] = n \exp\left(0 + \frac{\sigma^2 t^2}{2}\right)$$

where the last step is because X_i are iid and by Gaussian's moment generating function.

Thus take log on both sides we get

$$\mathbb{E}[Y] \leq \frac{\log n}{t} + \frac{\sigma^2 t}{2}$$

where to find minimum take $t = \dfrac{\sqrt{2\log n}}{\sigma}$ we get

$$\mathbb{E}[Y] \leq \sigma\sqrt{2\log n}.$$

But note if we were to compute a bound of $Z = \max\{X_1, -X_1, \dots, X_n, -X_n\}$ it suffices us to bound both the positive part and the negative part. But their bounds are the same since we're using the same moment generating function (mean is still 0). Thus this is the bound for what we want.

Now we check that for $\sigma = 2^{-n+1}$ we have

$$\mathbb{E}[Z] \leq \frac{\sqrt{2\log n}}{2^{n-1}} = o\left(\frac{n}{2^n}\right) = o\left(\frac{1}{n^2}\right) < \infty$$

and we are done (where the order sign really cares only when n large).

<u>Conclude the limit is a Brownian motion:</u>

By theorem in class, we only need to check 2 things in order to prove this: tightness of μ_{X_n} and weak convergence of finite dimensional distributions. By the first part of this question we are done with weak convergence of finite dimensional distributions.

But we have

$$\begin{aligned}\mathbb{P}\left(\lim_{n\to\infty} |X_n| \geq N\right) &\leq \mathbb{P}\left(\sum_{n=1}^{\infty} ||X_{n+1} - X_n||_\infty \geq N\right) \\ &= \mathbb{P}\left(\sum_{n=1}^{\infty} \max\left\{|Y_1|, \dots, |Y_l|\right\} \geq N\right) \to 0\end{aligned}$$

as $N \to \infty$ as right hand side is a convergent sequence, as is shown above. Hence they are a.s. Cauchy. Hence uniformly bounded.

Now, we only need to show uniform equicontinuity to show they are a tight family. To do so we note that at each point the local increment is the limit of a random walk with step size $N(0,1)$, since at each step the increment in slope is $\frac{N(0,2^{n-1})}{2^{n-1}} = 1$. But using the fact that the probability of the limit of a random walk goes to ∞ is 0, we get that with probability 1 the local increment is bounded at each point. So it's equicontinuous.

Moreover, since $X_n \in C[0,1]$ which is a family from a compact set to a metric space, hence equicontinuity implies uniformly equicontinuity. So we are done.

So we know the convergent holds.

$\square$