1. *Lévy's construction of Brownian motion.* This is nicely covered in RW, Chapter 1, section 6. The approach we took in lecture was different only in the way we expressed it. (The particular way we did the proof comes from Blumenthalsś book, *Excursions of Markov Processses.* Here, we want to emphasize only that the essential ingredients are (i) a few elementary facts about Gaussian random vectors, (ii) the bound on tail probabilities of standard (mean 0, variance 1) normal random variable Y,

$$\mathbb{P}(Y > x) \leq \frac{e^{-x^2/2}}{x\sqrt{2\pi}}, \tag{1}$$

(iii) the Borel-Cantelli Lemma, and (iv) the existence of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a countable collection of independent, standard normal random variables $\{Y_q\}_{q\in D}$, where D is the set of dyadic rationals in $[0,1]$.

The bound (1) follows by a simple integration by parts. The Borel-Cantelli Lemma says that if $\sum_1^\infty \mathbb{P}(A_i) < \infty$, then

$$\mathbb{P}\left(\cap_{n=1}^\infty \cup_{m=n}^\infty A_m\right) = 0.$$

The existence of a probability space supporting a countable, independent set of $N(0,1)$ random variables follows from a general result about constructing infinite products of probability spaces, and is also a consequence of a far more general result of Daniell-Kolmogorov, which will be treated below. We review next the facts about Gaussian (normal) random vectors, that are needed. These are really just the fact used in proving Corollary 2 of the previous lecture, but we provide a little more detail here. Given a non-negative definite $n \times n$-matrix C and an n-vector m, a random vector $Z = (Z_1, \ldots, Z_n)$ is said to have a joint normal (or Gaussian) distribution with mean m and covariance matrix C, its characteristic function is

$$\phi_Z(u) \stackrel{\triangle}{=} \mathbb{E}\left[e^{i\langle u,Z\rangle}\right] = e^{i\langle u,m\rangle-(1/2)\langle u,C\rangle}, \qquad u \in \mathbb{R}^n.$$

If C is positive definite, the joint density of $(Z_1, \ldots, Z_n)$ is

$$f(z) = \frac{1}{\sqrt{2\pi \det(C)}} e^{-(1/2)\langle z-m, C^{-1}(z-m)\rangle},$$

where $z = (z_1, \ldots, z_n)$ and $\langle\cdots,\cdot\rangle$ denotes the usual inner product. As the terminology suggests it is indeed the case that $m_i = \mathbb{E}[Z_i]$ and $C_{i,j} = \mathrm{Cov}(Z_i, Z_j)$ for all $1 \leq i,j \leq n$. Using the characteristic functions it is not hard to prove: (a) a vector of independent normal random variables is jointly normal; (b) the elements of a jointly normal random vector are mutually independent if and only if the covariance of any two of them is zero; (c) if Z is a jointly normal n-dimensional random vector and if A is an $m \times n$ matrix, AZ is jointly normal.

A stochastic process $\{X(t)\}_{t\in I}$, where I is a general index set, is said to be *Gaussian* if $(X(t_1), \ldots, X(t_n))$ is jointly normal for every set $\{t_1, \ldots, t_n\}$ of elements of I. The argument of Corollary 2 of the previous lecture show that if X is a Gaussian process and $\mathrm{Cov}(X(t), X(s)) = t \wedge s$, then X has independent increments. This argument applies to any sequence $(X(t_1), \ldots, X(t_n)$ of jointly normal random variables indexed by an increasing sequence of time: if $\mathrm{Cov}(X(t_i), X(t_j)) =$

$t_i \wedge t_j$¡ then the differences $X(t_1), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent. This is useful in proving Lévy's construction works.

The idea of L'evy's construction is to define a sequence of continuous, Gaussian processes $\{W_n(t)\,;\, 0 \le t \le 1\}$ so that W_n has the finite dimensional distributions of Brownian motion when the times are restricted to the set D_n of dyadic rationals of the form

$$D_n = \left\{ \frac{k}{2^n}\,;\, 0 \le k \le 2^n \right\},$$

Then one tries to obtain Brownian motion as a limit of $\{W_n\}$.

Now it is alway easy to build the process W_n. For example, if we let $\xi_1, \ldots \xi_{2^n}$ be a sequence of independent, $N(0, 2^{-n})$ random variables, then defining

$$W_n\left(\frac{k}{2^n}\right) = \sum_1^k \xi_i,$$

and linearly interpolating between these dyadic rationals, will do the trick, because then the increments of W are $W_n(k/2^n) - W((k-1)/2^n) = \xi_k$ and these are independent with the correct means and variances. However, if there is no probabilistic relation between W_n for different n, the only type of limit we can hope to extract is a weak limit, that is, limit in distribution. This can in fact be done but requires some sophisticated theory of weak convergence of probability measures on metric spaces. L'evy's idea was to build each new W_{n+1} from W_n by adding just the right Gaussian random variables to $W_n(t)$ at the points t in $D_{n+1} - D_n$ so that W_{n+1} will have the distribution of Brownian motion on D_{n+1}. When this construction is carried out, one discovers that it is easy to prove that the W_n converge a.s. in sup norm in the space of continuous functions on $[0, 1]$, and that the limit is Brownian motion.

The construction takes place on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, as discussed above in (iv) supporting a countable collection of independent $N(0, 1)$ random variables $\{Y_q\,;\, q \in D\}$, where $D = \cup_{n \ge 0} D_n$ is the set of all dyadic rationals. The processes W_n are recursively defined as follows. First, for $n = 0$, $W_0(0) = 0$ and $W_0(1) = y_1$, and $W_0(t)$ is defined at times $0 < t < 1$ by linear interpolation, so that $W_0(t) = tY_1$. Then, for each $n \ge 0$,

$$W_{n+1}(t) = \begin{cases} W_n(t), & t \in D_n \\ W_n(t) + \frac{1}{2^{(n+1)/2}} Y_t, & t \in D_{n+1} - D_n, \end{cases} \tag{2}$$

defines W_{n+1} on D_{n+1}, and it is extended to all of $[0, 1]$ by linear interpolation.

Notice that at each step, the random variables $\{Y_t\,;\, t \in D_{n+1} - D_n\}$ added to W_n to create W_{n+1} are independent of W_n. At each step, $W_n(\cdot)$ will be a linear combination of $\{Y_t\,;\, t \in D_n\}$ with time-dependent coefficients and hence will be Gaussian. Then it is easy to establish inductively that $\mathrm{Cov}(W_n(t), W_n(s)) = t \min s$ for $t, s \in D_n$, and hence will match the distributions defined in Corollary 1 of the previous lecture on the dyadic rationals. Notice also that $W_{n+1}(t) = W_n(t)$ on D_n, for all $n \ge 0$. Finally, observe that

$$\sup\{|W_{n+1}(t) - W_n(t)|;\, 0 \le t \le 1\} = \sup\{|Y_t|\,;\, t \in D_{n+1} - D_n\} \tag{3}$$

This is most easily seen by drawing an example of the relationship between W_n and W_{n+1}.

By (3) and the bound (1),

$$\begin{aligned}\mathbb{P}\left(\sup_{[0,1]}|W_{n+1}(t)-W_n(t)|\geq\frac{1}{n^2}\right) &\leq \sum_{t\in D_{n+1}-D_n}\mathbb{P}\left(|Y_t|\geq\frac{2^{(n+1)/2}}{n^2}\right)\\ &\leq 2^n\left[\frac{n^2}{2^{(n+1)/2}}\exp\{-2^n/n^4\}\right]\leq 2^{n/2}n^2\exp\{-2^n/n^4\}\end{aligned}$$

This is a summable sequence and so it follows by the Borel Cantelli lemma that with probability one, $\sum_1^\infty \sup_{[0,1]}|W_{n+1}(t)-W_n(t)|<\infty$, which implies that $\{W_n\}$ is Cauchy in sup norm. Therefore, almost surely there exists a continuous function $W(t)$, $0\leq t\leq 1$ such that $\sup_{[0,1]}|W_n(t)-W(t)|\to 0$ as $n\to\infty$. This W is Brownian motion.

The argument here is only sketched and the student should refer to RW for details,

2. *Construction of Brownian motion by way of the Daniell-Kolmogorov theorem.* This is covered in RW if you skip around. Here I will summarize the idea. It revolves around construction what is called, a canonical realization of a process from a specification of its finite-dimensional distributions. The motivating question is as follows. Suppose that for every finite set of times, $0\leq t_1<t_2<\cdots t_n$, we are given a probability measure $\mu_{t_1,\ldots,t_n}$ on $\mathbb{R}^n$. Under what conditions will there exist a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a stochastic process X on that space, such that for every positive integer n and every set of times $0\leq t_1<t_2<\cdots t_n$, the probability law of $(X(t_1),X(t_2),\ldots,X(t_n))$ is $\mu_{t_1,\ldots,t_n}$, that is,

$$\mathbb{P}\left((X(t_1),X(t_2),\ldots,X(t_n))\in U\right)\mu_{t_1,\ldots,t_n}(U)\qquad\text{for all Borel } U\subset\mathbb{R}^n?$$

This question frequently arises in constructing stochastic processes. One is interested in modeling a random phenomenon by a stochastic process. The modeling assumptions will determine what the finite dimensional distributions $\mu_{t_1,\ldots,t_n}$ should be and the mathematical problems is to construct a process fitting these distributions. Construction of a mathematical Brownian motion process is exactly a question of this type. We saw in the previous lecture notes that the simple assumptions that a process W have both independent, stationary increments and continuous paths, that $W(0)=0$, and that $W(1)$ have zero mean and unit variance, imply

$$\begin{aligned}&\mathbb{P}\left((W(t_1),W(t_2),\ldots,W(t_n))\in U\right)\\ &= \int\cdots\int \mathbf{1}_U(x_1,\ldots,x_n)\frac{1}{(2\pi)^{n/2}}\frac{\exp\{-\frac{x_1^2}{2t_1}-\frac{(x_2-x_1)^2}{2(t_2-t_1)}-\cdots-\frac{(x_n-x_{n-1})^2}{2(t_n-t_{n-1})}\}}{\sqrt{t_1}\sqrt{t_2-t_1}\cdots\sqrt{t_n-t_{n-1}}}\,\mathrm{d}x_1\ldots\mathrm{d}x_n,\end{aligned}\tag{4}$$

whenever $0<t_1<t_2<\cdots t_n$. The problem of showing mathematical Brownian motion exists is to construct a process with these finite dimensional distributions, which in addition has continuous paths.

Consider a family of probability measures

$$\mathcal{M}=\{\mu_{u_1,\ldots,u_n}\,;\,n\geq 1,\,0\leq u_1<u_2<\cdots u_n\}$$

To simplify notation, we let $U = \{u_1, \ldots, u_n\}$ and $\mu_U = \mu_{t_1,\ldots,t_n}$, we let $\mathbb{R}^U$ denote the set of functions from U to $\mathbb{R}$, and we think of the μ_U as a measure on $\mathbb{R}_U$. If $S \subset U$, let π_S^U denote the projection of $\mathbb{R}^U$ onto $\mathbb{R}^S$, defined by $\pi_S^U\left((x_u)_{u\in U}\right) = (x_s)_{s\in S}$.

Now suppose that there is in fact a stochastic process X whose finite-dimensional distributions are those specified by the family $\mathcal{M}$. Then, if $S \subset U$, for finite sets S and U,

$$\mu_S(B) = \mathbb{P}\left((X(s))_{s\in S} \in B\right) = \mathbb{P}\left(\pi_S^U\left[(X(u))_{u\in U}\right] \in B\right) = \mu_U\left((\pi_S^U)^{-1}(B)\right),$$

for all Borel sets B in $\mathbb{R}^S$. This is a simple consistency condition that the family $\mathcal{M}$ must satisfy in order to have an associated stochastic processes. The Daniell-Kolmogorov theorem says that it is also sufficient and provides a construction of the associated stochastic process. (The full-blown Daniell-Kolmogorov theorem actually considers a much more general situation: $\mathbb{R}$ is replaced by any Lusin space—a topological space homeomorphic to a Borel subset of a compact metric space, and the index set $[0,\infty)$ is replaced by a general index set. For simplicity, we state here only the special case directly relevant to our needs.)

The proof of the Daniell-Kolmogorov theorem constructs a so-called 'canonical' realization of a stochastic process associated to $\mathcal{M}$. For the underlying probability space Ω, it uses the space $\mathbb{R}^{[0,\infty)}$ of all paths from $[0,\infty)$ to $\mathbb{R}$; we denote an element of $\mathbb{R}$ by $\omega = (w(t))_{t\in[0,\infty)}$. If U is a subset of $[0,\infty)$, π_U denotes the project of $\mathbb{R}^{[0,\infty)}$ to $\mathbb{R}^U$: $\pi_U\left((w(t))_{t\in[0,\infty)}\right) = (\omega(u))_{u\in U}$. If U is finite and if B is a Borel subset of $\mathbb{R}^U$, considered as a finite dimensional vector space with the standard topology, a set of the form

$$(\pi_U)^{-1}(B)$$

is called a cylinder subset of $\mathbb{R}^{[0,\infty)}$. Let $\mathcal{C}$ denote the collection of all cylinder sets. It is not hard to see that $\mathcal{C}$ is an algebra of sets.

Finally, we define the canonical process X $\mathbb{R}^{[0,\infty)}$. This is simply the coordinate function process:

$$X(t)(\omega) = \omega(t), \qquad t \in [0,\infty).$$

This notation gives us all we need to state the special case of the Daniell-Kolmogorov theorem for $\mathbb{R}$-valued stochastic processes indexed by $[0,\infty)$.

Theorem 1 *Let $\mathcal{M} = \{\mu_U\,;\, |U| < \infty,\ U \subset [0,\infty),\ \mu_U$ is a prob. meas. on $\mathbb{R}^U\}$ be a set of finite dimensional probability measures. Suppose that*

$$\mu_S = \mu_U \circ (\pi_S^U)^{-1} \qquad \text{for any} \;\; S \subset U \subset [0,\infty) \;\; \text{where } |U| < \infty. \tag{5}$$

Then there exists a unique probability measure μ on $(\mathbb{R}^{[0,\infty)}, \sigma(\mathcal{C}))$ such that

$$\mu_U = \mu \circ (\pi_U)^{-1} \qquad \text{for every finite subset } U \text{ of } [0,\infty). \tag{6}$$

The canonical process on $(\mathbb{R}^{[0,\infty)}, \sigma(\mathcal{C}), \mu)$ is a stochastic process whose finite dimensional distributions are given by $\mathcal{M}$.

We sketch very succinctly the main idea of the proof. This is simply that, because of the compatibilty condition (5), the equation (6) consistently defines a finitely additive measure on the cylinder sets $\mathcal{C}$. One then uses the Caratheodory extension theorem to show it extends to a countably additive measure on $\sigma(\mathcal{C})$. The fact that the canonical process has the finite-dimensional distributions specified by $\mathcal{M}$ is a matter of chasing through definitions:

$$\mu\left(\{\omega\,;\,(X(u))_{u\in U}\in B\}\right)=\mu\left(\{\omega\,;\,(\omega(u))_{u\in U}\in B\}\right)=\mu\left((\pi_U)^{-1}(B)\right)=\mu_U(B).$$

Corollary 1 *There is a probability space $(\Omega,\mathcal{F},\mathbb{P})$ supporting a stochastic process Z such that $Z(0)=0$ and such that for every $0<t_1<t_2<\cdots t_n<\infty$, the joint density of $(Z(t_1),\ldots,Z(t_n))$ is*

$$p_n(x_1,\ldots,x_n)=\frac{1}{(2\pi)^{n/2}}\frac{\exp\{-\frac{x_1^2}{2t_1}-\frac{(x_2-x_1)^2}{2(t_2-t_1)}-\cdots-\frac{(x_n-x_{n-1})^2}{2(t_n-t_{n-1})}\}}{\sqrt{t_1}\sqrt{t_2-t_1}\cdots\sqrt{t_n-t_{n-1}}}.$$

To prove the corollary, it is only necessary to prove the compatibility of the finite-dimensional distributions defined by the given densities, and this is left as an exercise. The process Z then has the correct finite-dimensional distributions of a standard Brownian motion, by Corollary 1 of the the previous lecture notes. However it does not have continuous paths. So we may think of Z as a pre-Brownian motion.

How does one get a process with continuous paths from the pre-Brownian motion Z?. Given a process, X, we say that a process X' is a version of X if $\mathbb{P}(X(t)=X'(t))=1$ for all t. It turns out that one can construct a continuous version of Z. For this one uses a very nice criterion of Kolmogorov. It is worth stating this in some generality, and this is done in RW, pp. 59-60. Here we state only what we need for constructing Brownian motion:

Lemma 1 *Kolmogorov's Lemma (special case). Let $X=\{X(t);t\geq 0\}$ be a real-valued process. Assume that*

$$\mathbb{E}\left[\mid X(t)-X(s)\mid^{\alpha}\right]\leq C|t-s|^{1+\beta}. \tag{7}$$

for some constant C. Then there exists a version X' of X that is Hölder continuous of order θ for each $\theta<\beta/\alpha$.

It is easy to apply this to the pre-Brownian Z. Simply note that if Y is a normal random variable with mean 0 and variance σ^2,

$$\mathbb{E}\left[Z^{2k}\right]=\sigma^{2k}(2k-1)(2k-3)\cdots 1.$$

It is easy to obtain this by integration by parts and induction. For the pre-Brownian Z, $Z(t)-Z(s)$ is normal with mean 0 and variance $t-s$, when $t>s$. Thus

$$\mathbb{E}\left[|Z(t)-Z(s)\mid^{2k}\right]=(2k-1)(2k-3)\cdots 1|t-s|^k.$$

It follows that for every $k > 1$ there is a version of Z whose paths are almost-surely Hölder with constant θ for all $\theta < (k-1)/2k$. Since k can be arbitrarily large it follows there is a version of Z whose paths are Hölder continuous with order θ for all $\theta < 1/2$. Therefore we have obtain another construction of Brownian motion and, in addition, we have learned something about the Hölder continuity of its paths. It turns out the Brownian paths are a.s. not Hölder continuous of any order $\theta \geq 1/2$.