Third Assignment: Solutions

1. Since $P(X(p) > n) = (1-p)^n$, $n = 0, 1, 2, \ldots$, we have, for $x \geq 0$,

$$P(X(p)/p > x) = P(X(p) > [px] + 1) = (1-p)^{[px]+1} \to e^{-x}, \quad \text{as } p \to 0.$$

The conclusion follows from the fact that e^{-x} is continuous for $x > 0$.

2. (i) We only give sufficient conditions. First assume that the probability measure Q has density, with respect to the Lebesgue measure, f. Then

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\theta \cdot x} \Phi(\theta) d\theta.$$

Second, assume that $f(x)$ is positive definite:

$$\sum_{j=1}^{N} \sum_{k=1}^{N} c_j \bar{c}_k f(x_j - x_k) \geq 0,$$

for all $N \in \mathbb{N}$, all $c_1, \ldots, c_N \in \mathbb{C}$, and all $x_1, \ldots, x_N \in \mathbb{R}^d$. Then $\Phi(\theta) \geq 0$, by Bochner's theorem.
(ii) If R is a rotation and if X is a random variable with values in $\mathbb{R}^d$ and law Q, then $\Phi(\theta) = Ee^{i\theta \cdot X}$. Hence $\Phi(R\theta) = Ee^{i(R\theta)\cdot X} = Ee^{i\theta \cdot (R^T X)} = Ee^{i\theta \cdot X} = \Phi(\theta)$.
(iib) No: take, e.g., the probability measure Q on $\mathbb{R}^2$ with density

$$f(x, y) = c(x^2 + y^2) e^{-(x^2+y^2)}.$$

Then Q is invariant under rotations and has characteristic function

$$\Phi(\theta_1, \theta_2) = c\pi(1 - (\theta_1^2 + \theta_2^2)/4) e^{-(\theta_1^2+\theta_2^2)/4},$$

which takes both positive and negative values.

3. If there are two probability measures Q', Q'' corresponding to the same characteristic function, then the alternating sequence $Q', Q'', Q', Q'', \ldots$ satisfies the assumptions of Lévy's theorem and so it converges weakly to a probability measure. The only way for this sequence to converge is if $Q' = Q''$.

4. (i) Since $(X_1 + \cdots + X_n)/\sqrt{n}$ is a linear combination of $(X_1, \ldots, X_n)$ [which has normal law on $\mathbb{R}^n$], it follows immediately that $Z := (X_1 + \cdots + X_n)/\sqrt{n}$ has normal law on $\mathbb{R}^1$. That it is *standard* is clear from the fact that $E(Z) = 0$, $E(Z^2) = (1/n) \times n = 1$.
(ii) We have

$$\Phi(\theta) = \int_{-\infty}^{\infty} e^{i\theta x} \frac{1/\pi}{1 + x^2} dx = e^{-|\theta|},$$

as can be easily established by a contour integral. Therefore,

$$E[e^{i\theta X_j}] = e^{-|\theta|},$$

for all j. Let $W := (X_1 + \cdots + X_n)/n$. Then

$$E[e^{i\theta W}] = E \prod_{j=1}^{n} e^{i\theta X_j/n} = \prod_{j=1}^{n} E[e^{i(\theta/n)X_j}],$$

by independence, and so W has characteristic function

$$E[e^{i\theta W}] = \left(e^{-|\theta/n|}\right)^n = e^{-|\theta|},$$

which is identical to the characteristic function of X_1. Hence W and X_1 have the same law.
(iii) Indeed, the function

$$\Phi(\theta) := e^{-|\theta|^\alpha}$$

is positive semidefinite and hence a characteristic function of some probability measure Q_α. The stated property follows exactly in the same manner as above [part (i) corresponds to $\alpha = 2$, and part (ii) to $\alpha = 1$]. If $\alpha > 2$, the function above is not positive semidefinite.

5. (i) As in the hint, let U be uniform in $(-1, 1)$. Then, for all n,

$$U \stackrel{\text{(d)}}{=} \sum_{j=1}^{n} \frac{\xi_j}{2^j} + \frac{U}{2^n}.$$

The right-hand side converges weekly to $\sum_{j=1}^{\infty} \frac{\xi_j}{2^j}$. Hence the later has the law of U.
(ii) The left-hand side is the characteristic function of U. The characteristic function of $\sum_{j=1}^{\infty} \frac{\xi_j}{2^j}$ is the limit of the characteristic function of $\sum_{j=1}^{n} \frac{\xi_j}{2^j}$, as $n \to \infty$. The characteristic function of $\sum_{j=1}^{n} \frac{\xi_j}{2^j}$ is $\prod_{j=1}^{n} \cos(x/2^j)$. Hence $\lim_{n\to\infty} \prod_{j=1}^{n} \cos(x/2^j)$ exists and is equal to $(\sin x)/x$.
(iii) Set $x = \pi/2$. Then $(\sin x)/x = 2/\pi$. From the formula $\cos(2x) = 1 + 2\cos^2(x)$, it follows that

$$\cos(\pi/2^{j+1}) = \frac{\sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}}{2},$$

where the square root symbol appears j times. Hence

$$\frac{2}{\pi} = \prod_{j=1}^{\infty} \frac{\sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}^{\,(j \text{ times})}}{2}.$$

6. If $X(n)$, $n = 1, 2, \ldots$ is a tight family of random variables in $\mathbb{R}^d$ then it is obvious that, for each j, $X_j(n)$, $n = 1, 2, \ldots$ is a tight family of random variables in $\mathbb{R}$. For the converse, fix $\varepsilon > 0$ and pick $M_j > 0$ such that $P(|X_j(n)| > M_j) \le \varepsilon/d$, for all n. Let $M := \max(M_1, \ldots, M_d)$. Then $P(\max(|X_1(n)|, \ldots, |X(n)|) > M) \le \sum_{j=1}^{d} P(|X_j(n)| > M) \le \varepsilon$, for all n. Hence, for all n, the probability that $X(n)$ belongs to the compact set $K := [-1, 1]^d$ is at least $1 - \varepsilon$, for all n, and hence the sequence if tight.

7. Let $f : \mathbb{R}^d \to \mathbb{R}$ be continuous and bounded. Since g is continuous, it follows that $f \circ g : S \to \mathbb{R}$ is continuous and bounded. Therefore, $Q_n \circ (f \circ g)^{-1} \stackrel{\text{(d)}}{\longrightarrow} Q \circ (f \circ g)^{-1}$. But $Q_n \circ (f \circ g)^{-1} = (Q_n \circ g^{-1}) \circ f^{-1}$, and $Q \circ (f \circ g)^{-1} = (Q \circ g^{-1}) \circ f^{-1}$.

8. (i) If $\varphi \in C[0,1]$ then $\varphi(g(\varphi)) = 0$. Let $\tau := g(\varphi)$. If $\tau < 1$ and there is $\varepsilon > 0$ such that $\varphi > 0$ on $(\tau - \varepsilon, \tau)$ and $\varphi < 0$ on $(\tau, \tau + \varepsilon)$ then g is continuous at φ. To see this, let φ_n be a sequence of continuous functions such that $\varphi_n \to \varphi$, uniformly. Then φ_n will eventually be positive on $(\tau - \varepsilon, \tau)$ and negative on $(\tau, \tau + \varepsilon)$, implying that all the zeros of φ_n on $(\tau - \varepsilon, \tau + \varepsilon)$ converge to τ. So, necessarily, any discontinuity point φ of g must have the property that it has constant sign (or zero) on a small neighborhood of its last zero τ. Intuitively, a Brownian motion cannot do this, almost surely. The event that a Brownian motion B touches 0 at some point t_0 but is positive (respectively, negative) on a neighborhood of t_0 implies that the infimum (respectively, supremum) of the Brownian motion B on the same interval equals 0. By time scaling and space translation, it suffices to prove that $P(\sup_{0 \le t \le 1} B(t) = x) = 0$ for all x. This follows because $\sup_{0 \le t \le 1} B(t)$ is a random variable with continuous distribution function.

(ii) Since $S_n = 0$ if and only if n is even, it suffices to compute $P(T_{2m} = 2k)$. We have

$$\begin{aligned} P(T_{2m} = 2k) &= P(S_{2k} = 0, S_{2k+1} \neq 0, \ldots, S_{2m} \neq 0) \\ &= P(S_{2k} = 0) P(S_{2k+1} \neq 0, \ldots, S_{2m} \neq 0 \mid S_{2k} = 0) \\ &= P(S_{2k} = 0) P(S_1 \neq 0, \ldots, S_{2m-2k} \neq 0) \\ &= P(S_{2k} = 0) P(S_{2m-2k} = 0). \\ &= \binom{2k}{k} 2^{-2k} \binom{2m-2k}{m-k} 2^{-(2m-2k)} \\ &= \binom{2k}{k} \binom{2m-2k}{m-k} 2^{-2m}. \end{aligned}$$

(Alternatively, you can count the number of paths of length $2m$, starting from 0 and which hit 0 for the last time at time $2k$, and find that this is equal to the product of the binomial coefficients above.)

(iii) Let $Z_n(t) := B(nt)/\sqrt{n}$ if $t \in (1/n)\mathbb{Z}$, and let $Z_n(t)$ to be obtained by linear interpolation for other values of t. We know that the random sequence Z_n converges weakly to B. Since g is continuous at B, almost surely, we have that $g(Z_n)$ converges weakly to $T = g(B)$ and so

$$P(g(Z_n) \le x) \to P(g(B) \le x),$$

for all x such that $P(g(B) = x) = 0$. But

$$g(Z_n) = \sup\{t \le 1 : \ S_{[nt]} = 1\} = \frac{1}{n} T_n.$$

Therefore

$$P(T_n \le nx) \to P(T \le x), \qquad (*)$$

for all x such that $P(T = x) = 0$. From the formula for $P(T_{2m} = 2k)$ and Stirling's approximation we have

$$2mP(T_{2m} = 2k_m) \to \frac{2/\pi}{\sqrt{x(1-x)}} =: f(x), \quad \text{if } k_m/m \to x, \quad \text{as } m \to \infty.$$

In particular,

$$2mP(T_{2m} = 2[mx]) \to f(x), \quad \text{as } m \to \infty.$$

By the dominated convergence theorem,

$$\int_0^x 2mP(T_{2m} = 2[mt])dx \to \int_0^x f(t)dt, \quad \text{as } m \to \infty.$$

But the left-hand side equals $P(T_{2m} \leq 2mx)$ (here, we use the fact that T_{2m} takes only even values). Hence

$$P(T_{2m} \leq 2mx) \to \int_0^x f(t)dt. \qquad (**)$$

Comparing (*) and (**), we have

$$P(T \leq x) = \int_0^x f(t)dt = \frac{2}{\pi} \int_0^x \frac{dt}{\sqrt{t(1-t)}} = \frac{2}{\pi} \arcsin \sqrt{x}.$$

(Moreover, since $P(T = x) = 0$, for all x, convergence (*) holds for all x.)