1. Proof. (a) Let \( x \in [0,1] \) and \( \delta > 0 \), we want to show that there exist \( n \) such that \( n\lambda \mod 1 \in (x-\delta,x+\delta) \), for which it suffices to show that there exists \( n_0 \) and \( y \in (-\delta,\delta) \) such that \( n_0\lambda \equiv y \mod 1 \). Then take \( n \) to be an appropriate multiple of \( n_0 \).

To prove the reduced proposition, choose \( m \) such that \( 10^{-m} < \delta \). Note that there exist \( n_1 < n_2 \) such that
\[
(n_1\lambda \mod 1) \quad \text{and} \quad (n_2\lambda \mod 1)
\]
have the same first \( m \) digits after decimal point, since there are only finitely many possibilities for \( m \) digits. Letting \( n_0 = (n_2 - n_1) \cdot \text{sgn} \lambda \) completes the proof.

(b) We can choose an interval \( \delta \) until we find an interval with length less than \( |\delta| \), which is a contradiction. Hence one of the two halves satisfies \( |\delta|/2 \).

We claim that there exists \( k \) such that \( |A \cap J_k| > (1-\delta)|J_k| \). Otherwise, \( |A \cap J_k| \leq (1-\delta)|J_k| \) for all \( k \), then
\[
|A| \leq \sum_{n=1}^{\infty} |A \cap J_n| \leq (1-\delta) \sum_{n=1}^{\infty} |J_n| < (1-\delta) \cdot \frac{1}{1-\delta} |A| = |A|,
\]
contradiction.

Now we show that \( T \) is ergodic. Suppose \( A \) is an invariant set of positive measure. Given \( \epsilon \), from part (b) we can choose an interval \( J \) such that \( |A \cap J| \geq (1-\delta)|J| \). Furthermore, we can assume that \( |J| < \delta \).

Otherwise we can divide \( J \) into two intervals \( J_1 \) and \( J_2 \) with same length, if \( |A \cap J| < (1-\delta)|J| \) then
\[
|A \cap J| = |(A \cap J_1)| + |(A \cap J_2)| \leq (1-\delta)(|J_1| + |J_2|) = (1-\delta)|J|,
\]
which is a contradiction. Hence one of the two halves satisfies \( |A \cap J_i| \geq (1-\delta)|J_i| \). Continue this process until we find an interval with length less than \( \delta \).

From part (a) we know that there exist \( n_1 < n_2 < \cdots < n_k \) such that \( |T^{n_i}J| \) is a disjoint sequence and
\[
\sum_{i=1}^{k} |T^{n_i}J| = k|J| \geq 1 - \frac{\delta}{2}.
\]
To see this, suppose \( J = (a,b) \), we can find \( n_1 \) such that \( T^{n_1}a \in (0,\delta/2) \) and then find \( n_2 \) such that \( T^{n_2}a \in (T^{n_1}b,T^{n_1}b+\delta/4) \), in general, find \( n_i \) such that \( T^{n_i}a \in (T^{n_{i-1}}b,T^{n_{i-1}}b+\delta/2^i) \). Continue this process until we reach \( T^{n_k}b > 1 - \frac{\delta}{2} \), which will happen because \( T^{n_k}b - T^{n_{k-1}}b \geq |J| \). The ‘gaps’ left by \( T^{n_i}J \) has length at most \( \delta/2 + \cdots + \delta/2^k \) + \( \delta/2 \delta \leq \frac{5}{2}\delta \). Now,
\[
|A| \geq \sum_{i=1}^{k} |A \cap T^{n_i}J| = \sum_{i=1}^{k} |T^{-n_i}A \cap J| = \sum_{i=1}^{k} |A \cap J| \geq k(1-\delta)|J| \geq (1-\delta)(1-\frac{5}{2}\delta).
\]

Let \( \delta \to 0^+ \), we obtain that \( |A| = 1 \), which implies the ergodicity of \( T \).

2. Proof. (a) This is a direct consequence of Birkhoff’s Ergodic Theorem applied to \( \phi = \chi_A \).

(b) Suppose that \( y = x + \delta \) with \( |\delta| < \epsilon \). Then \( T^n x \equiv T^n y + \delta \mod 1 \). If \( T^n y \in [a+\epsilon,b-\epsilon] \), then \( T^n x \in [a + \epsilon - |\delta|, b - \epsilon + |\delta|] \subset [a,b] \), so \( N_n([a+\epsilon,b-\epsilon],y) \leq N_n([a,b],x) \). The other part can be shown similarly.

(c) Without loss of generality, we can assume that \( J \) is closed (since \( \lambda \) is irrational, we would meet each endpoint by at most once). Suppose that \( J = [a,b] \).
First, assume that \(0 < a < b < 1\). Choose \(M = \max\{a^{-1}, (1-b)^{-1}, 2(b-a)^{-1}\} + 1\). Denote \(E_A = \{y \in [0, 1] : \lim_{n \to \infty} N_n(A,y) = P(A)\}\). By part (a), both \(E_{(a+1/m,b-1/m)}\) and \(E_{(a-1/m,b+1/m)}\) have probability 1. Let
\[
E = \bigcap_{m \geq M} \left(E_{[a+\frac{1}{m}, b-\frac{1}{m}]} \cap E_{[a-\frac{1}{m}, b+\frac{1}{m}]}\right),
\]
then \(E\) has probability 1 and thus is dense in \([0, 1]\). Given \(x \in [0, 1]\), we can find \(y \in E\) such that \(|x-y| < \frac{1}{m}\) for all large \(m\). It follows from part (b) that
\[
\lim_{n \to \infty} N_n(A,x) \geq P(A) - \frac{2}{m}, \quad \lim_{n \to \infty} N_n(A,x) \leq P(A) + \frac{2}{m}.
\]
Let \(m \to \infty\), we obtain that
\[
\lim_{n \to \infty} N_n(A,x) = P(A).
\]
Now consider the case where \([0,b)\) \((b < 1)\). It can be proved similarly based on the following variation of part (b): For \(\epsilon < \frac{1}{2} \min\{b, 1-b\}\) and \(|x-y| < \epsilon\) it holds that
\[
N_n([\epsilon, b-\epsilon],y) \leq N_n([0,b],x) \leq N_n([0,b] \cup [1-\epsilon, 1],y).
\]
The case where \([a,1)\) \((a > 0)\) can be proved based on the following variation: For \(\epsilon < \frac{1}{2} \min\{a, 1-a\}\) and \(|x-y| < \epsilon\) it holds that
\[
N_n([a+\epsilon, 1-\epsilon],y) \leq N_n([a,1],x) \leq N_n([0 \cup [a,1],y).
\]
The last case \(J = [0,1]\) is trivial.

3. **Proof.** (a) First we show that \(\phi\) maps dyadic squares are mapped into rectangles of the same area. Suppose that \(I = [(k-1)/2^n, k/2^n] \times [(l-1)/2^n, l/2^n]\). Then
\[
\phi^{-1}(I) = \begin{cases} 
\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right] \times \left[\frac{l-1}{2^n}, \frac{l}{2^n}\right], & l \leq 2^{n-1} \\
\left[\frac{1}{2}, \frac{k-1}{2^n}, \frac{1}{2} + \frac{k}{2^n}\right] \times \left[\frac{2^n-l}{2^n} - 1, \frac{l}{2^n} - 1\right], & l > 2^{n-1}
\end{cases}
\]
It is easy to verify that \(m(I) = m(\phi^{-1}(I)) = 2^{-2n}\) in both cases. As dyadic squares generates the Borel algebra, the conclusion follows from the fact that the collection
\[
\{\text{Borel set } O : m(\phi^{-1}(O)) = m(O)\}
\]
is a \(\sigma\)-algebra.

(b) We claim that
\[
\lim_{n \to \infty} m(T^{-n}A \cap B) = m(A)m(B)
\]
for all measurable \(A, B\).

First we show (1) for dyadic rectangle \(A\) and \(B\). Suppose that \((\omega_1, \omega_2, \ldots)\) and \((\xi_1, \xi_2, \ldots)\) are binary representations of \(x, y \in [0,1]\). Then
\[
T : ((0,\omega_2,\ldots), (\xi_1, \xi_2, \ldots)) \mapsto ((\omega_2, \omega_3, \ldots), (0, \xi_1, \xi_2, \ldots))
\]
\[
((1,\omega_2,\ldots), (\xi_1, \xi_2, \ldots)) \mapsto ((\omega_2, \omega_3, \ldots), (1, \xi_1, \xi_2, \ldots)),
\]
which means that \(T\) shifts the first digit of the binary representation of \(x\) to the binary representation of \(y\), and thus \(T^{-1}\) does the reversed shift. A dyadic rectangle \(I\) can be characterized by fixed values on finitely many components in both binary representations of \(x\) and \(y\), and let \(x(I)\) and \(y(I)\) be the sets of indices of the finitely many components. For dyadic rectangles \(A\) and \(B\), we can always find \(N\) large enough such that \(\min x(T^{-N}A) > x(B)\) and \(y(T^{-N}A) = \emptyset\). Following a similar argument as in Problem 1 of Homework 1, we see that \(P(T^{-n}A \cap B) = 2^{-|x(T^{-n}I)|}2^{-|x(B)|} = P(A)P(B)\) for all \(n \geq N\). Hence (1) holds for dyadic rectangle \(A\) and \(B\).
Next we show (1) holds when $A$ and $B$ are open sets. Assume $A$ and $B$ are not empty. It is well-known that a non-empty open set can be written as a union of countably many non-overlapping dyadic rectangles (non-overlapping means that no two of them have a common interior point), hence we can write $A = \bigcup_{i=1}^{\infty} I_i$ and $B = \bigcup_{i=1}^{\infty} J_i$, where $I_i$’s and $J_i$’s are dyadic rectangles. Let $A_n = \bigcup_{i=1}^{n} I_i$ and $B_n = \bigcup_{i=1}^{n} J_i$, then $P(A_n) \to P(A)$ and $P(B_n) \to P(B)$. Note that

$$T^{-j} A \cap B = (T^{-j} A_n \cap B_n) \cup (T^{-j} A_n \cap (B \setminus B_n)) \cup (T^{-j} (A \setminus A_n) \cap B),$$

where the union is disjoint. Thus

$$\liminf_{j \to \infty} P(T^{-j} A \cap B) \geq \lim_{n \to \infty} (T^{-j} A_n \cap B_n) = P(A_n)P(B_n).$$

Let $n \to \infty$, we have that

$$\liminf_{j \to \infty} P(T^{-j} A \cap B) \geq P(A)P(B).$$

(2)

On the other hand,

$$\limsup_{j \to \infty} P(T^{-j} A \cap B) \leq \lim_{n \to \infty} P(T^{-j} A_n \cap B_n) + P(B \setminus B_n) + P(T^{-j} (A \setminus A_n))$$

$$= P(A_n)P(B_n) + P(B \setminus B_n) + P(A \setminus A_n).$$

Let $j \to \infty$, we have that

$$\limsup_{j \to \infty} P(T^{-j} A \cap B) \leq P(A)P(B),$$

which, combines with (2), yields (1).

Now, for measurable $A, B$ and $\epsilon > 0$, we can find open sets $O_1 \supset A$ and $O_2 \supset B$ such that $P(O_1 \setminus A) < \epsilon$ and $P(O_2 \setminus B) < \epsilon$. Similarly to the above, from

$$T^{-j} O_1 \cap O_2 = (T^{-j} A \cap B) \cup (T^{-j} A \cap (O_2 \setminus B)) \cup (T^{-j} (O_1 \setminus A) \cap O_2),$$

we can derive that

$$\limsup_{j \to \infty} P(T^{-j} A \cap B) \leq P(O_1)P(O_2) \leq (P(A) + \epsilon)(P(B) + \epsilon).$$

$$\liminf_{j \to \infty} P(T^{-j} A \cap B) \geq P(O_1)P(O_2) - P(O_1 \setminus A) - P(O_2 \setminus B) \geq P(A)P(B) - 2\epsilon.$$

Finally, let $\epsilon \to 0^+$ to complete the proof. \hfill \Box

4. Proof. (a)

$$a_m = k \iff k \cdot 10^n \leq 2^m < (k + 1) \cdot 10^n \text{ for some } n$$
$$\iff n + \log_{10} k \leq m \log_{10} 2 < n + \log_{10}(k + 1) \text{ for some } n$$
$$\iff m\lambda \mod 1 \in [\log_{10} k, \log_{10}(k + 1)) \quad (\lambda = \log_{10} 2).$$

(b) Let $J = [\log_{10} k, \log_{10}(k + 1))$ and $T : [0, 1] \to [0, 1]$ as $Tx = (x + \lambda) \mod 1$. Since $a_m = k$ iff $T^m(0) \in J$ and thus

$$N_n(J, \lambda) = \frac{1}{n} \sum_{m=0}^{n-1} \delta(k - a_m).$$

It follows from Problem 2(c) that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} \delta(k - a_m) = \lim_{n \to \infty} N_n(J, 0) = P(J) = \log_{10} \left(1 + \frac{1}{k}\right).$$
Proof. (a) We can show by induction that the graph of \( \omega \mapsto T^n \omega \) consists of \( 2^{n-1} \) tents, and each tent connects \( \left( \frac{k}{2^n}, 0 \right) \) and \( \left( \frac{2k-1}{2^n}, 1 \right) \) by a line segment and \( \left( \frac{2k-1}{2^n}, 1 \right) \) and \( \left( \frac{k}{2^n}, 0 \right) \) by another line segment.

The base case \( n = 1 \) is trivial. Suppose that the claim holds for some \( n \). According to the induction hypothesis,

\[
T^n(\omega) = 2^n \left( \omega - \frac{k - 1}{2^{n-1}} \right), \quad \omega \in \left[ \frac{k - 1}{2^{n-1}}, \frac{2k - 1}{2^n} \right].
\]

Hence

\[
T^{n+1}(\omega) = \begin{cases} 
2^{n+1} \left( \omega - \frac{k - 1}{2^{n+1}} \right), & \omega \in \left[ \frac{k - 1}{2^{n+1}}, \frac{4k - 3}{4n+1} \right] \\
2 \left[ 1 - 2^n \left( \omega - \frac{k - 1}{2^n} \right) \right], & \omega \in \left[ \frac{4k - 3}{4n+1}, \frac{2k - 1}{2^n} \right] \\
2^{n+1} \left( \omega - \frac{2k - 2}{2^{n+1}} \right), & \omega \in \left[ \frac{2k - 2}{2^{n+1}}, \frac{4k - 3}{4n+1} \right] \\
2^{n+1} \left( \omega - \frac{2k - 1}{2^{n+1}} \right), & \omega \in \left[ \frac{4k - 3}{4n+1}, \frac{2k - 1}{2^n} \right],
\end{cases}
\]

that is, there is a tent on \( \left[ \frac{2k - 2}{2^{n+1}}, \frac{2k - 1}{2^n} \right] \) of desired format. Similarly we can show that the graph of \( T^{n+1} \) is also a tent on \( \left[ \frac{2k - 1}{2^{n+1}}, \frac{2k}{2^n} \right] \). Each tent in the graph of \( T^n \) corresponds to two tents in the graph of \( T^{n+1} \), and there are \( 2^{n-1} \) tents in the graph \( T^n \), so there are \( 2^n \) tents in the graph of \( T^{n+1} \). The claim holds for \( n+1 \).

Now consider \( T^{−n}J_{r,k} \). It is clear from the graph of \( T^n \) that \( T^{−n}J_{r,k} \) consists of \( 2^{n-1} \) copies of two intervals \( I \) and \( J \), where \( I \) is \( T^{−n}J_{r,k} \) restricted on \( [0, 2^{−n}] \) and \( J \) is \( T^{−n}J_{r,k} \) restricted on \( [2^{−n}, 2^{−n-1}] \) (\( I \) and \( J \) are the inverse under a tent). In fact, \( J \) and \( I \) are symmetric about \( x = 2^{−n} \), and it is easy to see that \( I = I_{r,n+r} \), which is a dyadic interval of length \( 2^{−(n+r)} \). Therefore, \( T^{−n}J_{r,k} \) consists of union of \( 2^n \) non-overlapping dyadic intervals of length \( 2^{−n+r} \).

(b) From the graph of \( T^r \) it is clear that \( T^{−r}A \cap I_{r,k} \) is just a translation of \( T^{−r}J_{r,0} \) or \( T^{−r}A \cap I_{r,1} \). Note that \( T^{−r}A \cap I_{r,1} \) is just a reflection of \( T^{−r}A \cap I_{r,0} \) around \( x = 2^{−r} \), hence they have the same measure and \( m(T^{−r}A \cap I_{r,k}) = m(T^{−r}A \cap I_{r,0}) \). It is easy to see that

\[
T^{−r}A \cap I_{r,0} = \{2^{−r} \omega : \omega \in A\}.
\]

By scaling property of Lebesgue measure,

\[
T^{−r}A \cap I_{r,0} = 2^{−r}m(A),
\]

hence

\[
T^{−r}A \cap I_{r,k} = m(T^{−r}A \cap I_{r,0}) = 2^{−r}m(A) = m(I_{r,k})m(A).
\]

(c) Let \( r = 1 \) in part (b), we obtain that

\[
m \left( T^{−1}A \cap \left[ 0, \frac{1}{2} \right] \right) = m \left( T^{−1}A \cap \left[ \frac{1}{2}, 1 \right] \right) = \frac{1}{2} m(A)
\]

thus

\[
m(T^{−1}A) = m \left( T^{−1}A \cap \left[ 0, \frac{1}{2} \right] \right) + m \left( T^{−1}A \cap \left[ \frac{1}{2}, 1 \right] \right) = m(A),
\]

which implies that \( T \) is measure preserving. To show that \( T \) is strong mixing, we want to show that

\[
\lim_{n \to \infty} m(T^{−n}A \cap B) = m(A)m(B) \tag{3}
\]

for all measurable \( A \) and \( B \). First we show that (3) holds for all measurable \( A \) and dyadic interval \( B \). Suppose that \( |B| = 2^{−k} \). For \( n \geq k \), \( B \) can be written as a union of \( 2^{−(n-k)} \) non-overlapping dyadic intervals.
of length $2^{-n}$, say, $\{I_j\}$. Then
\[
m(T^{-n}A \cap B) = m \left( \bigcup_j T^{-n}A \cap I_j \right) = \sum_j m(T^{-n}A \cap I_j)\]
\[
= 2^{-(n-k)} \cdot m(A) 2^{-k} \quad \text{(by part (b))}
= 2^{-n} m(A) = m(A)m(B).
\]
It follows (3) for all measurable $A$ and dyadic interval $B$. Then we can extend $B$ to open sets and then to finally general measurable sets, as what we have done in Problem 3(b).

6. Proof. (a) It is clear that $\hat{T}(\hat{\Omega}) = \hat{\Omega}$. Let $A \in \hat{\mathcal{F}}$, then
\[
P(\hat{T}^{-1}A) = \int_{\hat{T}^{-1}A} \frac{dx}{\pi \sqrt{x(1-x)}}
= 2 \int_A \frac{1}{\pi \sqrt{1-y}} dy \quad \text{(let } y = Tx)\]
\[
= \int_A \frac{dy}{\pi \sqrt{y(1-y)}} = \hat{P}(A).
\]
(b) It is clear that $S$ is bijective. We show that $S$ is measure-preserving. In fact, let $A \in \hat{\mathcal{F}}$, then
\[
P(S^{-1}A) = \int_{S^{-1}A} dx = \int_A \frac{dy}{\pi \sqrt{y(1-y)}} \quad \text{(let } y = Sx)\]
\[
= \hat{P}(A).
\]
(c) Direct calculation gives that
\[
ST(x) = \sin^2 \pi x,
\]
\[
\hat{T}S(x) = \hat{T} \left( \sin^2 \frac{\pi x}{2} \right) = 4 \sin^2 \frac{\pi x}{2} \cos^2 \frac{\pi x}{2} = \sin^2 \pi x,
\]
hence $ST = \hat{T}S$, or, $\hat{T} = STS^{-1}$, and thus $\hat{T}^{-n} = ST^{-n}S^{-1}$. For $A, B \in \hat{\mathcal{F}}$
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \hat{P}(\hat{T}^{-n}A \cap B) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \hat{P}(S(T^{-n}S^{-1}A \cap S^{-1}B))
= \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} P(T^{-n}S^{-1}A \cap S^{-1}B)
= P(S^{-1}A)P(S^{-1}B) \quad \text{(T is ergodic by Problem 5(b))}
= \hat{P}(A)\hat{P}(B),
\]
which implies that $\hat{T}$ is ergodic. \qed