1. (a) Since $EX^2 < \infty$, $MSE(a) = E(X-a)^2 = EX^2 - 2aEX + a^2$. The first-order condition (FOC) is
\[
\frac{dMSE(a)}{da} = -2EX + 2a = 0,
\]
and the second-order sufficient condition (SOSC) is
\[
\frac{d^2MSE(a)}{da^2} = 2 > 0.
\]
Thus, $a = EX$ minimizes the MSE and the MSE evaluated at $a = EX$ is the variance of $X$, i.e.,
\[
MSE(EX) = E(X - EX)^2 = Var(X).
\]
(b) Assume that $X$ has a pdf $f(x)$ and cdf $F(x)$. Then
\[
MAD(a) = E|X-a| = \int_{-\infty}^{\infty} |x-a|f(x)dx
\]
\[
= \int_{-\infty}^{a} (a-x)f(x)dx + \int_{a}^{\infty} (x-a)f(x)dx
\]
\[
= a\int_{-\infty}^{a} f(x)dx - \int_{-\infty}^{a} xf(x)dx + \int_{a}^{\infty} xf(x)dx - a\int_{a}^{\infty} f(x)dx
\]
\[
= aF(a) - a(1-F(a)) - 2\int_{-\infty}^{a} xf(x)dx + \int_{-\infty}^{\infty} xf(x)dx.
\]
Let $G(x) = \int xf(x)dx$. Since the second moment (and thus the first moment) exists, $\lim_{x \to -\infty} G(x)$ and $\lim_{x \to \infty} G(x)$ exist. Therefore,
\[
\frac{dMAD(a)}{da} = F(a) + af(a) - (1-F(a)) + af(a) - 2af(a) = 2F(a) - 1 = 0,
\]
and $a$ is the median of $X$ (SOSC is $2f(a) > 0$).

2. Fix $x_0$ and take a decreasing sequence $\{x_n\}$ such that $\lim_{n \to \infty} x_n = x_0$. Let $C_n = \{\omega : X(\omega) \leq x_n\}$. Then $C_n$ is a decreasing sequence of events and we have $\lim_{n \to \infty} C_n = \cap_{n=1}^{\infty} C_n = \{\omega : X(\omega) \leq x_0\}$. Therefore,
\[
\lim_{x \downarrow x_0} F_X(x) = \lim_{n \to \infty} F_X(x_n) = \lim_{n \to \infty} P(C_n) = P(\lim_{n \to \infty} C_n) = P(\{\omega : X(\omega) \leq x_0\}) = F_X(x_0).
\]
3.

We say that a set \( C \) is a Borel set if \( C \) can be obtained by a countable number of operations, starting from open sets, each operation consisting in taking unions, intersections, or complements (Rudin, 1976).

Define \( C_1 = [0, 1/3] \cup [2/3, 1], \) \( C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1], \) and so on. Then \( C_n \) is the union of \( 2^n \) disjoint closed intervals with length \( 3^{-n} \). Now the Cantor set is \( C = \cap_{n=1}^\infty C_n. \)

(a) Since \( C_n \) is the union of \( 2^n \) disjoint closed intervals (each is the complement of an open interval), \( C_n \) is a Borel set. So \( C = \cap_{n=1}^\infty C_n \) is a Borel set.

(b) Countable additivity: \( \mu(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n) \) if \( A_i \cap A_j = 0 \) for \( i \neq j \).

\[
\mu(C) = 1 - \mu(C^c) \\
= 1 - \mu(\cup_{n=1}^\infty C_n^c) \quad \because \text{De Morgan’s law} \\
= 1 - \sum_{n=1}^\infty \mu(C_n^c) \quad \because \text{countable additivity} \\
= 1 - \sum_{n=1}^\infty \left(\frac{2}{3}\right)^n \frac{1}{2} = 0.
\]

(c) Omitted.

4.

(a) Suppose that for some positive integer \( k > k_0 \), \( X \) has finite \( k \)th moment. Then \( E|X^k| < \infty \) and

\[
E|X^k| = \int_{\mathbb{R}} |x^k| f(x) dx = \int_{|x|>1} |x^k| f(x) dx + \int_{|x|\leq1} |x^k| f(x) dx \\
\geq \int_{|x|>1} |x^{k_0}| f(x) dx - \int_{|x|\leq1} |x^k| f(x) dx \\
\geq \int_{\mathbb{R}} |x^{k_0}| f(x) dx - \int_{|x|\leq1} |x^{k_0}| f(x) dx \quad \because |x^k| \leq |x^{k_0}| \\
= \infty,
\]

which contradicts to the assumption. Thus, \( X \) does not have finite \( k \)th moment for any \( k > k_0 \).

(b) Suppose that \( M_X(t) \) is well-defined on a neighborhood of 0 and is differentiable with respect to \( t \) (differentiation under the integral sign). Since \( e^{tX} \) is infinitely differentiable with respect to \( t \), by Taylor’s theorem,

\[
Ee^{tX} = M_X(0) + \frac{M_X^{(1)}(0)}{1!} t + \frac{M_X^{(2)}(0)}{2!} t^2 + \cdots + \frac{M_X^{(k_0)}(0)}{k_0!} t + R(t) \\
= 1 + EX \cdot t + \frac{EX^2}{2!} t^2 + \cdots + \frac{EX^{k_0}}{k_0!} t^{k_0} + R(t),
\]

where \( R(t) = o(|t|^{k_0}) \) is a remainder term. Since \( EX^{k_0} \) is not well-defined by assumption (\( \because E|X^{k_0}| = \infty \)), the right-hand side of the above equation is not well-defined. This contradicts to the assumption. Thus, \( M_X(t) \) is not well-defined on a neighborhood of 0.
5. 

Since $e^{tx} > 0$ for all $x$, by Markov’s inequality,

$$P(e^{tx} \geq e^{at}) \leq e^{-at}M_X(t),$$

for $-h < t < h$. Now note that $e^{tx} \geq e^{at}$ if and only if $x \geq a$ for $0 < t < h$ and $e^{tx} \geq e^{at}$ if and only if $x \leq a$ for $-h < t < 0$. This completes the proof.

6. 

Note that $P(X \geq 3) = \left(\frac{2}{3}\right)^3$. Then,

$$P(X = x|X \geq 3) = \frac{P(X = x, X \geq 3)}{P(X \geq 3)} = \frac{1}{3} \left(\frac{2}{3}\right)^{x-3},$$

for $x = 3, 4, 5, \ldots$, and zero elsewhere.

7. 

Since $X$ is a continuous random variable, $P(c < X < d) = P(c < X \leq d) = P(X \leq d) - P(X \leq c)$ and $P(X < c) = P(X \leq c)$. Using the table in the appendix, we find $c = 0.831$ and $d = 12.833$.

8. 

(a) Assume $x > 0$ and $y > 0$. $P(X > x) = 1 - P(X \leq x) = 1 - \int_{-\infty}^{x} \lambda e^{-\lambda t} dt = e^{-\lambda x}$ for a positive $\lambda$. Then

$$P(X > x + y|X > x) = \frac{P(X > x + y, X > x)}{P(X > x)} = \frac{P(X > x + y)}{P(X > x)} = \frac{P(X > x + y)}{e^{-\lambda x}} = e^{-\lambda y} = P(X > y).$$

(b) Since property (3.3.7) holds for $Y$, $P(Y > x + y) = P(Y > x)P(Y > y)$ for $x, y > 0$. Let $g(y) = 1 - F_Y(y)$. Then $g(x + y) = g(x)g(y)$ and $\log g(x + y) = \log g(x) + \log g(y)$. Since $Y$ is a continuous random variable, $g(\cdot)$ is differentiable. By differentiating $\log g(x + y) = \log g(x) + \log g(y)$ with respect to $x$ on both sides,

$$\frac{g'(x + y)}{g(x + y)} = \frac{g'(x)}{g(x)}.$$

By letting $x = 0$,

$$\frac{g'(y)}{g(y)} = \frac{g'(0)}{g(0)} = g'(0),$$
because \( g(0) = 1 - F_Y(0) = 1 \). Let \( g'(0) = -\lambda \) and solve the differential equation

\[
\frac{g'(y)}{g(y)} = \frac{d \log g(y)}{dy} = -\lambda.
\]

Then \( \log g(y) = -\lambda y + C_1 \) and \( g(y) = C_2 e^{-\lambda y} \) for some constants \( C_1 \) and \( C_2 \). Since we know \( g(0) = 1, C_2 = 1 \). Therefore, \( g(y) = 1 - F_Y(y) = e^{-\lambda y} \) or \( F_Y(y) = 1 - e^{-\lambda y} \) for \( y > 0 \).

9.

The random variable \( X \) is \( N(3,4^2) \). Thus,

\[
P(-1 < X < 9) = P\left(-1 < \frac{X - 3}{4} < 1.5\right)
= P(-1 < Z < 1.5)
= \Phi(1.5) - \Phi(-1) = 0.7745,
\]

where \( Z \) is the standard normal random variable and \( \Phi(\cdot) \) is its cdf.