

Probability and Statistics

Quiz 2 - Solutions

Question 1

Problem Statement:

Suppose $Z = X + U$, where X is an exponential random variable with parameter λ and U is a Uniform(0,1) random variable. Derive the first two moments of Z using MGF. (You have to derive the MGF of X and U first.) Assume X and U are independent with respect to any other random variable.

Note: 0.5 marks will be deducted if the correct domain of t is not written.

Solution

Step 1: MGF of X and U

For $X \sim \text{Exponential}(\lambda)$:

$$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

For $U \sim \text{Uniform}(0, 1)$:

$$M_U(t) = E[e^{tU}] = \int_0^1 e^{tu} du = \frac{e^t - 1}{t}, \quad t \neq 0, \quad M_U(0) = 1.$$

Step 2: MGF of Z

Since X and U are independent,

$$M_Z(t) = M_X(t) M_U(t) = \frac{\lambda}{\lambda - t} \cdot \frac{e^t - 1}{t}.$$

Define

$$A(t) = M_X(t) = \frac{\lambda}{\lambda - t}, \quad B(t) = M_U(t) = \frac{e^t - 1}{t},$$

so that $M_Z(t) = A(t)B(t)$.

Step 3: First Moment using MGF

Compute derivatives of $A(t)$ and $B(t)$:

$$A'(t) = \frac{\lambda}{(\lambda - t)^2}, \quad A''(t) = \frac{2\lambda}{(\lambda - t)^3}.$$

Hence,

$$A(0) = 1, \quad A'(0) = \frac{1}{\lambda}, \quad A''(0) = \frac{2}{\lambda^2}.$$

Now expand $B(t)$ using the Taylor series of e^t :

$$e^t = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \dots \Rightarrow \frac{e^t - 1}{t} = 1 + \frac{t}{2} + \frac{t^2}{6} + \dots$$

Thus,

$$B(0) = 1, \quad B'(0) = \frac{1}{2}, \quad B''(0) = \frac{1}{3}.$$

Now,

$$M'_Z(t) = A'(t)B(t) + A(t)B'(t).$$

At $t = 0$,

$$M'_Z(0) = A'(0)B(0) + A(0)B'(0) = \frac{1}{\lambda} + \frac{1}{2}.$$

Hence,

$$E[Z] = M'_Z(0) = \frac{1}{\lambda} + \frac{1}{2}.$$

Step 4: Second Moment using MGF

Differentiate again:

$$M''_Z(t) = A''(t)B(t) + 2A'(t)B'(t) + A(t)B''(t).$$

At $t = 0$:

$$\begin{aligned} M''_Z(0) &= A''(0)B(0) + 2A'(0)B'(0) + A(0)B''(0) \\ &= \frac{2}{\lambda^2} + 2\left(\frac{1}{\lambda}\right)\left(\frac{1}{2}\right) + \frac{1}{3} \\ &= \frac{2}{\lambda^2} + \frac{1}{\lambda} + \frac{1}{3}. \end{aligned}$$

Thus,

$$E[Z^2] = M''_Z(0) = \frac{2}{\lambda^2} + \frac{1}{\lambda} + \frac{1}{3}.$$

Final Answers:

$$E[Z] = \frac{1}{\lambda} + \frac{1}{2}, \quad E[Z^2] = \frac{2}{\lambda^2} + \frac{1}{\lambda} + \frac{1}{3}.$$

Note: Any other valid and correct derivation using the MGF of Z will be marked accordingly.

Question 2

Problem Statement:

Let X_1, X_2, X_3, \dots be a sequence of Laplacian (double-exponential) random variables with density

$$f_{X_n}(x) = \frac{n}{2} e^{-n|x|}, \quad x \in \mathbb{R}.$$

Show that $X_n \xrightarrow{P} 0$.

Proof 1: By Definition

We must show that for every $\varepsilon > 0$,

$$\mathbb{P}(|X_n| > \varepsilon) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Use the density to compute the tail probability. For some arbitrary $\varepsilon > 0$ we have,

$$\mathbb{P}(|X_n| > \varepsilon) = \int_{|x|>\varepsilon} \frac{n}{2} e^{-n|x|} dx = 2 \int_{\varepsilon}^{\infty} \frac{n}{2} e^{-nx} dx = n \int_{\varepsilon}^{\infty} e^{-nx} dx.$$

Now compute

$$n \int_{\varepsilon}^{\infty} e^{-nx} dx = n \left[\frac{-1}{n} e^{-nx} \right]_{x=\varepsilon}^{\infty} = e^{-n\varepsilon}.$$

Hence for any fixed $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} e^{-n\varepsilon} = 0.$$

Therefore $\mathbb{P}(|X_n| > \varepsilon) \rightarrow 0$, so by definition $X_n \xrightarrow{P} 0$.

Proof 2: Using Markov's Inequality

Markov's inequality states that for any nonnegative random variable Y and any $\varepsilon > 0$,

$$\mathbb{P}(Y > \varepsilon) \leq \frac{\mathbb{E}[Y]}{\varepsilon}.$$

Since X_n is not nonnegative in general, apply the inequality to $Y := |X_n|$. Hence

$$\mathbb{P}(|X_n| > \varepsilon) \leq \frac{\mathbb{E}[|X_n|]}{\varepsilon}.$$

Therefore,

$$\mathbb{E}[|X_n|] = \int_{-\infty}^{\infty} |x| \frac{n}{2} e^{-n|x|} dx = 2 \cdot \frac{n}{2} \int_0^{\infty} x e^{-nx} dx = n \int_0^{\infty} x e^{-nx} dx.$$

Evaluate the integral (use $\int_0^{\infty} x e^{-nx} dx = 1/n^2$):

$$\mathbb{E}[|X_n|] = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

Hence by Markov,

$$\mathbb{P}(|X_n| > \varepsilon) \leq \frac{\mathbb{E}[|X_n|]}{\varepsilon} = \frac{1}{n\varepsilon} \rightarrow 0.$$

Thus $X_n \xrightarrow{P} 0$.

Proof 3: Using Chebyshev's Inequality

Chebyshev's inequality: $\mathbb{P}(|X_n - \mathbb{E}[X_n]| > \varepsilon) \leq \text{Var}(X_n)/\varepsilon^2$.

Since f_{X_n} is even, the integrand $x f_{X_n}(x)$ is odd, so

$$\mathbb{E}[X_n] = \int_{-\infty}^{\infty} x f_{X_n}(x) dx = 0.$$

For the second moment:

$$\mathbb{E}[X_n^2] = \int_{-\infty}^{\infty} x^2 \frac{n}{2} e^{-n|x|} dx = 2 \cdot \frac{n}{2} \int_0^{\infty} x^2 e^{-nx} dx = n \int_0^{\infty} x^2 e^{-nx} dx.$$

Use $\int_0^{\infty} x^2 e^{-nx} dx = \frac{2}{n^3}$ (gamma function property) to get

$$\mathbb{E}[X_n^2] = n \cdot \frac{2}{n^3} = \frac{2}{n^2}.$$

Thus $\text{Var}(X_n) = 2/n^2$. By Chebyshev,

$$\mathbb{P}(|X_n| > \varepsilon) \leq \frac{\text{Var}(X_n)}{\varepsilon^2} = \frac{2}{n^2 \varepsilon^2} \rightarrow 0,$$

so again $X_n \xrightarrow{P} 0$.

Question 3

Problem Statement:

Give a procedure to convert samples from an Exponential random variable to Uniform(0,1). Justify why the procedure is correct.

Solution

We first recall the probability density function (pdf) and cumulative distribution function (cdf) of both distributions.

- **Uniform(0,1):**

$$f_U(u) = \begin{cases} 1, & 0 \leq u \leq 1, \\ 0, & \text{otherwise,} \end{cases} \quad F_U(u) = \begin{cases} 0, & u < 0, \\ u, & 0 \leq u \leq 1, \\ 1, & u > 1. \end{cases}$$

- **Exponential(λ):**

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$$

Part (a): Procedure

Property :

We can use the property **Universality of the Uniform**, which states:

For any continuous random variable X with cdf F_X , the random variable $F_X(X) \sim \text{Uniform}(0, 1)$.

Steps :

1. Draw a sample x from the exponential distribution $X \sim \text{Exp}(\lambda)$.
2. Compute $u = F_X(x) = 1 - e^{-\lambda x}$.
3. Return u . This gives a sample from $\text{Uniform}(0, 1)$.

Part (b): Justification

Claim: If $X \sim \text{Exp}(\lambda)$ and $U = F_X(X) = 1 - e^{-\lambda X}$, then $U \sim \text{Uniform}(0, 1)$.

Proof:

$$\begin{aligned} P(U \leq u) &= P(F_X(X) \leq u) \\ &= P(1 - e^{-\lambda X} \leq u) \\ &= P(e^{-\lambda X} \geq 1 - u) \\ &= P(-\lambda X \geq \ln(1 - u)) \\ &= P\left(X \leq -\frac{1}{\lambda} \ln(1 - u)\right) \\ &= F_X\left(-\frac{1}{\lambda} \ln(1 - u)\right) \\ &= 1 - \exp\left(-\lambda \cdot \left(-\frac{1}{\lambda} \ln(1 - u)\right)\right) \\ &= 1 - (1 - u) = u. \end{aligned}$$

Hence, $P(U \leq u) = u$ for all $u \in [0, 1]$.

Therefore, $U \sim \text{Uniform}(0, 1)$.

Note:

Methods such as the **Accept–Reject method** or any other approach that relies on Uniform random variables **cannot be used here**, because we are only given access to Exponential random variables. We cannot use a Uniform random variable to generate itself. **Such methods will receive 0 marks.**

Question 4

Problem Statement:

Given the joint probability density function

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} 1, & 0 < x_1 < 1, 0 < x_2 < 1, \\ 0, & \text{otherwise,} \end{cases}$$

and the transformation

$$\begin{cases} Y_1 = X_1 - X_2, \\ Y_2 = X_1 + X_2. \end{cases}$$

Find the joint probability density function $f_{Y_1, Y_2}(y_1, y_2)$.

Solution

Finding the Inverse Transformation

From the given transformation, we solve for X_1 and X_2 in terms of Y_1 and Y_2 :

$$\begin{aligned} Y_1 + Y_2 &= (X_1 - X_2) + (X_1 + X_2) = 2X_1 \quad \Rightarrow \quad X_1 = \frac{Y_1 + Y_2}{2}, \\ Y_2 - Y_1 &= (X_1 + X_2) - (X_1 - X_2) = 2X_2 \quad \Rightarrow \quad X_2 = \frac{Y_2 - Y_1}{2}. \end{aligned}$$

Therefore, the inverse mapping is:

$$\boxed{x_1 = \frac{y_1 + y_2}{2}, \quad x_2 = \frac{y_2 - y_1}{2}}$$

Calculating the Determinant of the Jacobian

The Jacobian matrix of the inverse transformation is:

$$J = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Computing the determinant:

$$\det(J) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Therefore, $\boxed{|J| = \frac{1}{2}}.$

Applying the Change of Variables Formula

By the transformation theorem for random variables:

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}\left(\frac{y_1 + y_2}{2}, \frac{y_2 - y_1}{2}\right) \cdot |J|.$$

Since $f_{X_1, X_2}(x_1, x_2) = 1$ on the unit square, we have:

$$f_{Y_1, Y_2}(y_1, y_2) = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

whenever the point (y_1, y_2) maps to a point in the unit square.

Determining the Support

The support of (Y_1, Y_2) consists of all points (y_1, y_2) such that:

$$0 < \frac{y_1 + y_2}{2} < 1 \quad \text{and} \quad 0 < \frac{y_2 - y_1}{2} < 1.$$

Multiplying through by 2:

$$\begin{cases} 0 < y_1 + y_2 < 2, \\ 0 < y_2 - y_1 < 2. \end{cases}$$

From these inequalities, we can derive:

$$\begin{aligned} y_2 - y_1 > 0 &\Rightarrow y_1 < y_2, \\ y_1 + y_2 > 0 &\Rightarrow y_1 > -y_2, \\ y_2 - y_1 < 2 &\Rightarrow y_1 > y_2 - 2, \\ y_1 + y_2 < 2 &\Rightarrow y_1 < 2 - y_2. \end{aligned}$$

Combining these constraints:

$$\max(-y_2, y_2 - 2) < y_1 < \min(y_2, 2 - y_2).$$

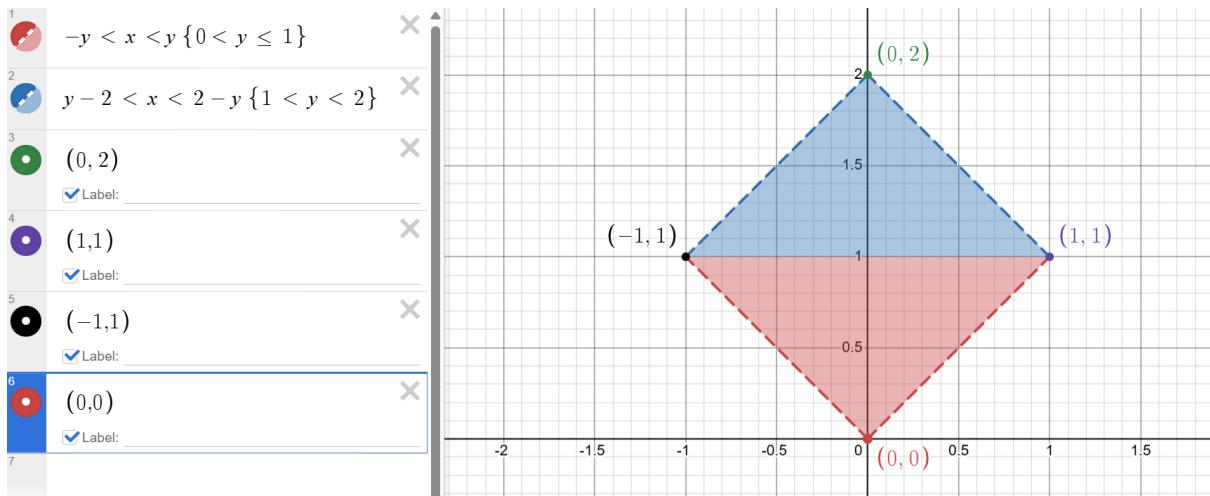
We can split this into two cases based on the value of y_2 :

- **Case 1:** If $0 < y_2 \leq 1$, then $-y_2 \geq y_2 - 2$ and $y_2 \leq 2 - y_2$, so:

$$-y_2 < y_1 < y_2.$$

- **Case 2:** If $1 < y_2 < 2$, then $y_2 - 2 > -y_2$ and $2 - y_2 < y_2$, so:

$$y_2 - 2 < y_1 < 2 - y_2.$$



Final Joint Density

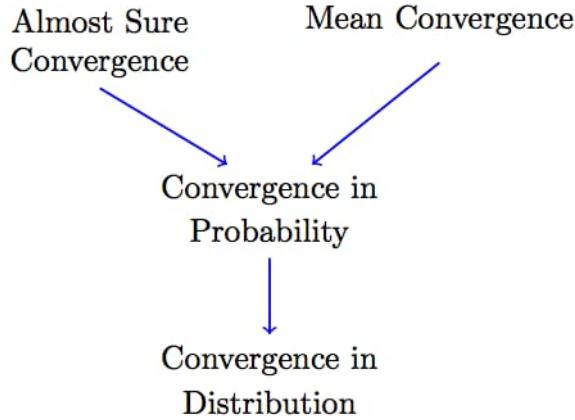
The joint probability density function of (Y_1, Y_2) is:

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} \frac{1}{2}, & 0 < y_2 \leq 1, -y_2 < y_1 < y_2, \\ \frac{1}{2}, & 1 < y_2 < 2, y_2 - 2 < y_1 < 2 - y_2, \\ 0, & \text{otherwise.} \end{cases}$$

The shaded region demonstrates the probability density in the (y_1, y_2) space. The X-axis represents y_1 and Y-axis represents y_2 .

Question 1b

2.5 marks for each part. No marks for correct answer without justification.



partial marks awarded for c,d if this relation has been used directly

a. Convergence in Distribution

For $y \geq 0$:

$$\begin{aligned} F_{Y_n}(y) &= P(Y_n \leq y) \\ &= P(nX_n \leq y) \\ &= P(X_n \leq y/n) \end{aligned}$$

We now use the CDF of X_n , $F_{X_n}(x) = 1 - e^{-nx}$, with $x = y/n$:

$$\begin{aligned} F_{Y_n}(y) &= F_{X_n}(y/n) \\ &= 1 - e^{-n(y/n)} \\ &= 1 - e^{-y} \end{aligned}$$

$$\lim_{n \rightarrow \infty} F_{Y_n}(y) = \lim_{n \rightarrow \infty} (1 - e^{-y}) = 1 - e^{-y} = F_Y(y)$$

Since the CDFs are identical for all n , convergence in distribution is shown.

b. Convergence in Probability

For convergence in probability ($Y_n \xrightarrow{P} Y$), we would need $\lim_{n \rightarrow \infty} P(|Y_n - Y| > \epsilon) = 0$ for all $\epsilon > 0$.

From part (1), we know $Y_n \sim \text{Exp}(1)$ for all n . We are given $Y \sim \text{Exp}(1)$. We assume Y_n and Y are independent random variables.

Let's compute the probability $P(|Y_n - Y| > \epsilon)$.

$$P(|Y_n - Y| > \epsilon) = P(Y_n - Y > \epsilon) + P(Y - Y_n > \epsilon)$$

We will compute each term by conditioning on one of the variables.

$$\begin{aligned}
 P(Y_n > Y + \epsilon) &= \int_0^\infty P(Y_n > y + \epsilon \mid Y = y) f_Y(y) dy \\
 &= \int_0^\infty P(Y_n > y + \epsilon) f_Y(y) dy \quad (\text{by independence}) \\
 &= \int_0^\infty (e^{-(y+\epsilon)}) \cdot (e^{-y}) dy \quad (\text{since } P(Y_n > x) = e^{-x} \text{ and } f_Y(y) = e^{-y}) \\
 &= \int_0^\infty e^{-y} e^{-\epsilon} e^{-y} dy \\
 &= e^{-\epsilon} \int_0^\infty e^{-2y} dy \\
 &= e^{-\epsilon} \left[\frac{e^{-2y}}{-2} \right]_0^\infty \\
 &= e^{-\epsilon} \left(0 - \left(\frac{e^0}{-2} \right) \right) = \frac{1}{2} e^{-\epsilon}
 \end{aligned}$$

The calculation is identical due to symmetry:

$$\begin{aligned}
 P(Y > Y_n + \epsilon) &= \int_0^\infty P(Y > x + \epsilon \mid Y_n = x) f_{Y_n}(x) dx \\
 &= \int_0^\infty P(Y > x + \epsilon) f_{Y_n}(x) dx \quad (\text{by independence}) \\
 &= \int_0^\infty (e^{-(x+\epsilon)}) \cdot (e^{-x}) dx \quad (\text{since } P(Y > x) = e^{-x} \text{ and } f_{Y_n}(x) = e^{-x}) \\
 &= e^{-\epsilon} \int_0^\infty e^{-2x} dx = \frac{1}{2} e^{-\epsilon}
 \end{aligned}$$

$$P(|Y_n - Y| > \epsilon) = \frac{1}{2} e^{-\epsilon} + \frac{1}{2} e^{-\epsilon} = e^{-\epsilon}$$

This probability is a positive constant that does not depend on n .

$$\lim_{n \rightarrow \infty} P(|Y_n - Y| > \epsilon) = \lim_{n \rightarrow \infty} e^{-\epsilon} = e^{-\epsilon}$$

Since $e^{-\epsilon} \neq 0$ for any finite $\epsilon > 0$, the condition for convergence in probability is not met.

c. Almost Sure Convergence

Theorem 7.5 (Borel Cantelli Lemma) can not be used to prove

$$Y_n \xrightarrow{a.s.} Y$$

since it is sufficient but not necessary

Using Theorem 7.6 from probabilitycourse.com (was mentioned as HW in slides).

Borel Cantelli Lemma

Self-Study: Theorem 7.5 (probabilitycourse.com)

Consider a sequence of random variables X_1, X_2, \dots . If for all ϵ we have

$$\sum_{n=1}^{\infty} P(|X_n - X| > \epsilon) < \infty$$

then $X_n \rightarrow X$ a.s.

- ▶ This is only a sufficient condition for almost sure convergence!
- ▶ Thm 7.6 (HW) gives necessary and sufficient conditions.

Theorem 7.6

Consider the sequence X_1, X_2, X_3, \dots . For any $\epsilon > 0$, define the set of events

$$A_m = \{|X_n - X| < \epsilon, \text{ for all } n \geq m\}.$$

Then $X_n \xrightarrow{\text{a.s.}} X$ if and only if for any $\epsilon > 0$, we have

$$\lim_{m \rightarrow \infty} P(A_m) = 1.$$

$$\begin{aligned} P(A_m) &= P\left(\bigcap_{n=m}^{\infty} \{|Y_n - Y| < \epsilon\}\right) \\ &= \prod_{n=m}^{\infty} P(|Y_n - Y| < \epsilon) \quad (\text{independence}) \\ &= \prod_{n=m}^{\infty} (1 - P(|Y_n - Y| > \epsilon)) \end{aligned}$$

From part (2), we know $P(|Y_n - Y| > \epsilon) = e^{-\epsilon}$.

$$P(A_m) = \prod_{n=m}^{\infty} (1 - e^{-\epsilon})$$

$$\implies P(A_m) = 0$$

Now, we take the limit:

$$\lim_{m \rightarrow \infty} P(A_m) = \lim_{m \rightarrow \infty} 0 = 0$$

Since $0 \neq 1$, the condition for Theorem 7.6 is not met.

$$\implies Y_n \not\stackrel{q.s.}{\rightarrow} Y$$

d. Convergence in Mean (L^1)

Convergence in L^1 requires $\lim_{n \rightarrow \infty} E[|Y_n - Y|] = 0$.

Let $Z = Y_n - Y$. We first derive the PDF of Z , $f_Z(z)$, by finding its CDF, $F_Z(z) = P(Z \leq z)$.

- **Case 1:** $z \geq 0$

$$F_Z(z) = P(Z \leq z) = P(Y_n - Y \leq z) = 1 - P(Y_n - Y > z).$$

From our calculation in part (2) (with z in place of ϵ), we know $P(Y_n - Y > z) = \frac{1}{2}e^{-z}$.

$$F_Z(z) = 1 - \frac{1}{2}e^{-z}.$$

- **Case 2:** $z < 0$

$$F_Z(z) = P(Z \leq z) = P(Y_n - Y \leq z) = P(Y - Y_n \geq -z).$$

Let $\epsilon = -z$. Since $z < 0$, $\epsilon > 0$.

$$F_Z(z) = P(Y - Y_n \geq \epsilon).$$

From our calculation in part (2), substituting $\epsilon = -z$, we get $F_Z(z) = \frac{1}{2}e^{-(z)} = \frac{1}{2}e^z$.

- For $z > 0$: $f_Z(z) = \frac{d}{dz}(1 - \frac{1}{2}e^{-z}) = \frac{1}{2}e^{-z}$.
- For $z < 0$: $f_Z(z) = \frac{d}{dz}(\frac{1}{2}e^z) = \frac{1}{2}e^z$.

$$f_Z(z) = \frac{1}{2}e^{-|z|}$$

Now we can compute $E[|Z|]$:

$$\begin{aligned} E[|Y_n - Y|] &= E[|Z|] = \int_{-\infty}^{\infty} |z| f_Z(z) dz \\ &= \int_{-\infty}^{\infty} |z| \frac{1}{2}e^{-|z|} dz \\ &= 2 \int_0^{\infty} z \cdot \frac{1}{2}e^{-z} dz \quad (\text{by symmetry}) \\ &= \int_0^{\infty} ze^{-z} dz \end{aligned}$$

This integral is the definition of the expected value of an $\text{Exp}(1)$ random variable, which is 1.

Since $E[|Y_n - Y|] = 1$ for all n , the limit is:

$$\lim_{n \rightarrow \infty} E[|Y_n - Y|] = \lim_{n \rightarrow \infty} 1 = 1$$

Because the limit is $1 \neq 0$, Y_n does not converge in L^1 to Y .