

# Homework 5

Week 6

6733172621 Patthadon Phengpinij

*Collaborators.* ChatGPT (for L<sup>A</sup>T<sub>E</sub>X styling and grammar checking)

## z-table

z	0.0	+0.1	+0.2	+0.3	+0.4	+0.5	+0.6	+0.7	+0.8	+0.9
0.0	0.5000	0.5398	0.5793	0.6179	0.6554	0.6915	0.7258	0.7580	0.7881	0.8159
1.0	0.8413	0.8643	0.8849	0.9032	0.9192	0.9332	0.9452	0.9554	0.9641	0.9713
2.0	0.9773	0.9821	0.9861	0.9893	0.9918	0.9938	0.9953	0.9965	0.9974	0.9981
3.0	0.9987									

Table 1: Standard Normal Distribution Table (z-table)

## 1 Week 6: Continuous Random Variables

**Problem 1.** For each statement below, determine if it is True or False.

- The value of a Probability Density Function (PDF),  $f(x)$ , can be greater than 1.
- The value of a Cumulative Distribution Function (CDF),  $F(x)$ , can be greater than 1.
- The integral of a valid PDF,  $\int_{-\infty}^{\infty} f(x) dx$ , over its entire range must equal 1.
- For any continuous random variable  $X$ , the probability  $P(X = c)$  is always 0.

### Solution.

- True.** A PDF can take values greater than 1, as long as the total area under the curve equals 1.
- False.** A CDF must always be between 0 and 1, inclusive.
- True.** The integral of a valid PDF over its entire range must equal 1, representing the total probability.
- True.** For continuous random variables, the probability of taking any specific value is 0, since there are infinitely many possible values.

**Problem 2.** A continuous random variable  $X$  has the following piecewise PDF:

$$f(x) = \begin{cases} cx & 0 \leq x \leq 2 \\ c(4-x) & 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

- a) Find the value of  $c$  that makes  $f(x)$  a valid PDF.  
 b) Sketch the graph of the PDF.

**Solution.** For a valid PDF, the total area under the curve must equal 1. We can find  $c$  by integrating  $f(x)$  over its entire range and setting the integral equal to 1.

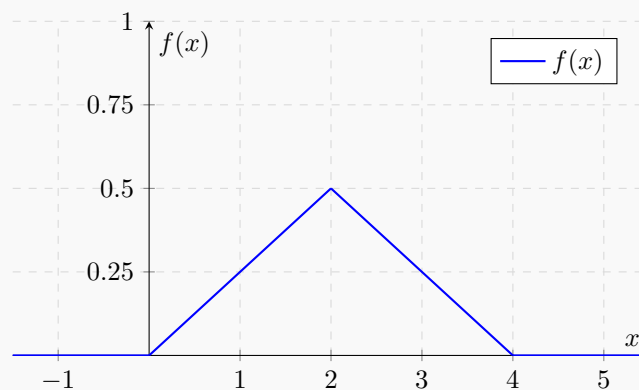
$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x) dx = \int_0^2 cx dx + \int_2^4 c(4-x) dx \\ &= c \left[ \frac{x^2}{2} \right]_0^2 + c \left[ 4x - \frac{x^2}{2} \right]_2^4 \\ &= c \left[ \frac{2^2}{2} - 0 + \left( 4(4) - \frac{4^2}{2} \right) - \left( 4(2) - \frac{2^2}{2} \right) \right] \\ &= c[2 + (16 - 8) - (8 - 2)] \\ 1 &= c[2 + 8 - 6] = 4c \end{aligned}$$

Thus, the value of  $c$  that would make this PDF valid is:  $c = \frac{1}{4}$  a).

This PDF would be:

$$f(x) = \begin{cases} \frac{1}{4}x & 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

- b). The graph of the PDF is as follows:



**Problem 3.** Using the PDF from Problem 2:

- Calculate the probability  $P(X > 2.5)$ .
- Derive the Cumulative Distribution Function (CDF),  $F(x)$ .

**Solution.** From the previous problem, we have the PDF:

$$f(x) = \begin{cases} \frac{1}{4}x & 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

To calculate  $P(X > 2.5)$ , we integrate the PDF from 2.5 to 4:

$$\begin{aligned} P(X > 2.5) &= \int_{2.5}^{\infty} f(x) dx \\ &= \int_{2.5}^4 f(x) dx + \int_4^{\infty} f(x) dx \\ &= \int_{2.5}^4 \frac{1}{4}(4-x) dx + \int_4^{\infty} 0 dx \\ &= \frac{1}{4} \left[ 4x - \frac{x^2}{2} \right]_{2.5}^4 + 0 \\ &= \frac{1}{4} [(16 - 8) - (10 - 3.125)] \\ &= \frac{1}{4} [8 - 6.875] \\ P(X > 2.5) &= \frac{1}{4} \cdot 1.125 = 0.28125 \end{aligned}$$

Thus,  $P(X > 2.5) = \boxed{0.28125}$  **a).**

To derive the CDF,  $F(x)$ , we integrate the PDF from the lower limit to  $x$ . Considering the piecewise nature of the PDF, we have:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{4}x & 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & 2 \leq x \leq 4 \\ 0 & x \geq 4 \end{cases}$$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \frac{1}{4}t dt & 0 \leq x \leq 2 \\ \int_0^2 \frac{1}{4}t dt + \int_2^x \frac{1}{4}(4-t) dt & 2 \leq x \leq 4 \\ 1 & x \geq 4 \end{cases}$$

Calculating the integrals:

$$\begin{aligned} \int_0^x \frac{1}{4}t dt &= \frac{1}{4} \cdot \left[ \frac{t^2}{2} \right]_0^x \\ &= \frac{1}{4} \cdot \frac{x^2}{2} \\ \int_0^x \frac{1}{4}t dt &= \frac{x^2}{8} \end{aligned}$$

And:

$$\begin{aligned}
 \int_2^x \frac{1}{4}(4-t) dt &= \frac{1}{4} \cdot \left[ 4t - \frac{t^2}{2} \right]_2^x \\
 &= \frac{1}{4} \cdot \left( (4x - \frac{x^2}{2}) - (8 - 2) \right) \\
 &= \frac{1}{4} \cdot (4x - \frac{x^2}{2} - 6) \\
 \int_2^x \frac{1}{4}(4-t) dt &= x - \frac{x^2}{8} - \frac{3}{2}
 \end{aligned}$$

Thus, the CDF  $F(x)$  is:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{8} & 0 \leq x \leq 2 \\ \left( \frac{x^2}{8} \right)_{x=2} + \left( x - \frac{x^2}{8} - \frac{3}{2} \right) & 2 \leq x \leq 4 \\ 1 & x \geq 4 \end{cases}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{8} & 0 \leq x < 2 \\ x - \frac{x^2}{8} - \frac{3}{2} & 2 \leq x < 4 \\ 1 & x \geq 4 \end{cases} \text{ b).}$$

**Problem 4.** Using the CDF from Problem 3, find the median of the random variable  $X$ .

**Solution.** From the previous problem, we have the CDF:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x^2}{8} & 0 \leq x \leq 2 \\ x - \frac{3}{2} & 2 \leq x \leq 4 \\ 1 & x \geq 4 \end{cases}$$

For a random variable  $X$ , the median is the value  $m$  such that  $F(m) = 0.5$ . To find the median, we need to solve for  $m$  in the equation  $F(m) = 0.5$ . We consider the piecewise nature of the CDF:

1. For  $0 \leq m \leq 2$ :

$$F(m) = \frac{m^2}{8} = 0.5$$

Solving for  $m$ :

$$m^2 = 4 \implies m = 2 \implies m \in [0, 2], \text{ valid.}$$

2. For  $2 \leq m \leq 4$ :

$$F(m) = m - \frac{3}{2} = 0.5$$

Solving for  $m$ :

$$m = 2 \implies m \in [2, 4], \text{ valid.}$$

Thus, the median of the random variable  $X$  is  $\boxed{2}$ .

**Problem 5.** A shuttle bus arrives at a student dormitory at a random time within a 15-minute window, from 8:00 AM to 8:15 AM. Let  $T$  be the student's waiting time in minutes if they arrive at exactly 8:00 AM.

- What is the distribution of  $T$ ? (Specify type and parameters).
- What is the probability that a student waits for more than 10 minutes?

**Solution. a).** The waiting time  $T$  follows a uniform distribution over the interval  $[0, 15]$  minutes. Thus, we can denote this as:  $T \sim \text{Uniform}(0, 15)$

**b).** To find the probability that a student waits for more than 10 minutes:  $P(T > 10)$ . The PDF of a uniform distribution over the interval  $[a, b]$  is given by:

$$f_T(t) = \begin{cases} \frac{1}{b-a} & a \leq t \leq b \\ 0 & \text{otherwise} \end{cases}$$

For our case,  $a = 0$  and  $b = 15$ , so the PDF is:

$$f_T(t) = \begin{cases} \frac{1}{15} & 0 \leq t \leq 15 \\ 0 & \text{otherwise} \end{cases}$$

To find  $P(T > 10)$ , we can integrate the PDF from 10 to 15:

$$P(T > 10) = \int_{10}^{15} f_T(t) dt = \int_{10}^{15} \frac{1}{15} dt = \frac{1}{15} \cdot (15 - 10) = \frac{5}{15} = \frac{1}{3}$$

Thus, the probability that a student waits for more than 10 minutes is  $\boxed{\frac{1}{3}}$ .

**Problem 6.** For the shuttle bus scenario in Problem 5:

- What is the expected waiting time,  $E[T]$ ?
- What is the variance of the waiting time,  $\text{Var}(T)$ ?

**Solution.** From Problem 5, we know that  $T \sim \text{Uniform}(0, 15)$ .

For a uniform distribution  $\text{Uniform}(a, b)$ :

- The expected value  $E[X] = \frac{a+b}{2}$
- The variance  $\text{Var}(X) = \frac{(b-a)^2}{12}$

The expected waiting time  $E[T]$  is:

$$E[T] = \frac{0 + 15}{2} = \frac{15}{2} = \boxed{7.5 \text{ minutes}} \text{ a).}$$

The variance of the waiting time  $\text{Var}(T)$  is:

$$\text{Var}(T) = \frac{(15 - 0)^2}{12} = \frac{225}{12} = \boxed{18.75 \text{ minutes}^2} \text{ b).}$$

## TO SUBMIT

**Problem 7.** For the shuttle bus scenario in Problem 5, what is the 80<sup>th</sup> percentile of the waiting time? In other words, find the time  $t$  (in minutes) such that 80% of waiting times are less than  $t$ .

**Solution.** From Problem 5, we know that  $T \sim \text{Uniform}(0, 15)$ .  
The CDF of a uniform distribution  $\text{Uniform}(a, b)$  is given by:

$$F_T(t) = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a} & a \leq t \leq b \\ 1 & t > b \end{cases}$$

For our case,  $a = 0$  and  $b = 15$ , so the CDF is:

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \frac{t}{15} & 0 \leq t \leq 15 \\ 1 & t > 15 \end{cases}$$

To find the 80th percentile, we need to solve for  $t$  in the equation  $F_T(t) = 0.8$ :

$$\frac{t}{15} = 0.8$$

Solving for  $t$ :

$$t = 0.8 \times 15 = 12$$

Thus, the 80th percentile of the waiting time is 12 minutes.

**Problem 8.** The lifespan of a particular brand of hard drive, in years, follows an exponential distribution with a mean of 4 years.

- a) What is the value of the parameter  $\lambda$  for this distribution?
- b) What is the probability that a new hard drive lasts for less than 1 year?

**Solution.** For an exponential distribution, the mean  $\mu$  is given by:

$$\mu = \frac{1}{\lambda}$$

Given that the mean lifespan of the hard drive is 4 years, we can find  $\lambda$ :

$$4 = \frac{1}{\lambda} \implies \lambda = \left[ \frac{1}{4} \right] \text{ a).}$$

The PDF of an exponential distribution is given by:

$$f(x) = \lambda e^{-\lambda x} \text{ for } x \geq 0$$

To find the probability that a new hard drive lasts for less than 1 year, we calculate:

$$P(X < 1) = \int_0^1 f(x) dx = \int_0^1 \frac{1}{4} e^{-\frac{1}{4}x} dx = \left[ -e^{-\frac{1}{4}x} \right]_0^1 = -e^{-\frac{1}{4}} + e^0 = 1 - e^{-\frac{1}{4}} \approx 0.2212$$

Thus, the probability that a new hard drive lasts for less than 1 year is  $1 - e^{-\frac{1}{4}}$  b).

### TO SUBMIT

**Problem 9.** For the hard drive in Problem 8, you discover an old one that is still working after 3 years. What is the probability that it will continue to work for at least another 4 years?

**Solution.** From Problem 8, we know that the lifespan of the hard drive follows an exponential distribution with parameter  $\lambda = \frac{1}{4}$ .

The exponential distribution has the memoryless property, which means that the probability of surviving an additional time period is independent of how long it has already survived.

Thus, we need to find  $P(X > 7 \mid X > 3)$ , which is equivalent to  $P(X > 4)$  due to the memoryless property.

$$P(X > 4) = e^{-\lambda \cdot 4} = e^{-\frac{1}{4} \cdot 4} = e^{-1} \approx 0.3679$$

Thus, the probability that the hard drive will continue to work for at least another 4 years is  $\boxed{e^{-1}}$ .

**Problem 10.** What is the **median** lifespan of the hard drives from Problem 8? Is this value smaller or larger than the mean lifespan, and why does this make intuitive sense for this distribution?

**Solution.** From Problem 8, we know that the lifespan of the hard drive follows an exponential distribution with parameter  $\lambda = \frac{1}{4}$ .

The CDF of an exponential distribution is given by:

$$F(x) = 1 - e^{-\lambda x} \text{ for } x \geq 0$$

To find the median lifespan, we need to solve for  $m$  in the equation  $F(m) = 0.5$ :

$$0.5 = 1 - e^{-\frac{1}{4}m}$$

$$e^{-\frac{1}{4}m} = 0.5$$

$$-\frac{1}{4}m = \ln(0.5)$$

$$m = -4 \ln(0.5)$$

$$m = 4 \ln(2)$$

Thus, the median lifespan of the hard drives is  $\boxed{4 \ln(2) \approx 2.7726 \text{ years}}$ .

### TO SUBMIT

**Problem 11.** The final exam scores for a large statistics class are normally distributed with a mean of 72 and a standard deviation of 12.

- a) What is the probability that a randomly selected student scores between 65 and 85?
- b) What is the probability a student scores exactly 80.00?

**Solution.** From the z-table provided, we can use the standard normal distribution to find the required probabilities.

STANDARD NORMAL DISTRIBUTION										
Z Score Table										
Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189

Negative Z Score Table
Positive Z Score Table

Let  $X$  be the random variable representing the exam scores, where  $X \sim N(72, 12^2)$ .

a). To find the probability that a randomly selected student scores between 65 and 85, we need to calculate  $P(65 < X < 85)$ . We first convert the scores to z-scores:

$$z_1 = \frac{65 - 72}{12} = -\frac{7}{12} \approx -0.5833$$

$$z_2 = \frac{85 - 72}{12} = \frac{13}{12} \approx 1.0833$$

Using the z-table, we find the probabilities corresponding to these z-scores:

$$P(Z < -0.58) \approx 0.2810$$

$$P(Z < 1.08) \approx 0.8599$$

Thus, the probability that a student scores between 65 and 85 is:

$$P(65 < X < 85) = P(Z < 1.08) - P(Z < -0.58) \approx 0.8599 - 0.2810 = 0.5789$$

So,  $P(65 < X < 85) \approx \boxed{0.5789}$  a).

b). For a continuous random variable, the probability of scoring exactly any specific value is 0. Thus,  $P(X = 80.00) = \boxed{0}$  b).



**Problem 12.** For the exam scores in Problem 11, the professor wants to give an ‘A’ grade to the top 10% of students. What is the minimum score a student must get to receive an ‘A’?

**Solution.** From Problem 11, we know that the exam scores follow a normal distribution:

$$X \sim N(72, 12^2)$$

To find the minimum score a student must get to receive an ‘A’, we need to find the score corresponding to the 90th percentile of the distribution. We first find the z-score that corresponds to the 90th percentile using the z-table:

$$z \approx 1.28$$

We then convert this z-score back to the original score using the formula:

$$X = \mu + z \cdot \sigma = 72 + 1.28 \cdot 12 \approx 72 + 15.36 = 87.36$$

Thus, the minimum score a student must get to receive an ‘A’ is approximately 87.36.

**Problem 13.** Given  $Z$  is a standard normal random variable, use a Z-table to find:

- a)  $P(Z < -1.4)$
- b)  $P(-0.5 < Z < 1.9)$

**Solution.** a). From the z-table, we find:

$$P(Z < -1.4) \approx \text{0.0808}$$

b). To find  $P(-0.5 < Z < 1.9)$ , we can use the z-table to find the probabilities:

$$P(Z < -0.5) \approx 0.3085 \text{ and } P(Z < 1.9) \approx 0.9713$$

Thus, the probability is:

$$P(-0.5 < Z < 1.9) = P(Z < 1.9) - P(Z < -0.5) \approx 0.9713 - 0.3085 = \text{0.6628}$$

**Problem 14.** Find the value of  $c$  for a standard normal variable  $Z$  such that:

- a)  $P(Z < c) = 0.99$
- b)  $P(Z > c) = 0.75$

**Solution.** a). From the z-table, we find:

$$P(Z < c) = 0.99 \Rightarrow c \approx \text{2.3}$$

b). To find  $P(Z > c) = 0.75$ , we can use the z-table to find the corresponding z-score:

$$P(Z < c) = 1 - P(Z > c) = 1 - 0.75 = 0.25 \Rightarrow c \approx \text{-0.7}$$

**Problem 15.** The weight of coffee in a “large” cup from a machine is  $X \sim N(400, 100)$  (in ml). The weight of milk added is  $Y \sim N(50, 25)$ . Assume  $X$  and  $Y$  are independent.

- a) Let  $T = X + Y$  be the total volume. What are  $E[T]$  and  $\text{Var}(T)$ ?  
 b) What is the probability the total volume exceeds 460 ml?

**Solution.** a). For independent random variables  $X$  and  $Y$ :

$$E[T] = E[X] + E[Y] = 400 + 50 = \boxed{450 \text{ ml}}$$

$$\text{Var}(T) = \text{Var}(X) + \text{Var}(Y) = 100 + 25 = \boxed{125 \text{ ml}^2}$$

b). To find the probability that the total volume exceeds 460 ml, we first find the distribution of  $T$ :

$$T \sim N(450, 125)$$

We need to calculate  $P(T > 460)$ . We first convert this to a z-score:

$$z = \frac{460 - 450}{\sqrt{125}} = \frac{10}{\sqrt{125}} \approx 0.9$$

Using the z-table, we find:

$$P(Z < 0.9) \approx 0.8159$$

Thus, the probability that the total volume exceeds 460 ml is:

$$P(T > 460) = 1 - P(Z < 0.9) \approx 1 - 0.8159 = \boxed{0.1841}$$

**Problem 16.**

**Solution.** Let  $D = B - C$ . Since  $B$  and  $C$  are independent normal random variables,  $D$  is also normally distributed:

$$D \sim N(E[B] - E[C], \text{Var}(B) + \text{Var}(C))$$

Calculating the mean and variance of  $D$ :

$$E[D] = 120 - 125 = -5$$

$$\text{Var}(D) = 16 + 9 = 25$$

Thus,  $D \sim N(-5, 25)$ .

We need to find  $P(D > 0)$ , which is equivalent to finding  $P(B > C)$ . We first convert this to a z-score:

$$z = \frac{0 - (-5)}{\sqrt{25}} = \frac{5}{5} = 1$$

Using the z-table, we find:

$$P(Z < 1) \approx 0.8413$$

Thus, the probability that the blueberry muffin is heavier than the chocolate chip muffin is:

$$P(D > 0) = 1 - P(Z < 1) \approx 1 - 0.8413 = \boxed{0.1587}$$

**Problem 17.** Let  $X$  be a random variable representing the outcome of a fair six-sided die roll. Let  $Y$  be a continuous random variable from a  $U(1, 6)$  distribution.

- Explain the fundamental difference between  $P(X = 4)$  and  $P(Y = 4)$ .
- Calculate  $P(X > 4)$  and  $P(Y > 4)$ .

**Solution. a).** The fundamental difference between  $P(X = 4)$  and  $P(Y = 4)$  lies in the nature of the random variables.

$X$  is a discrete random variable, meaning it can take on specific, countable values (in this case, the integers 1 through 6). Therefore,  $P(X = 4)$  represents the probability of rolling a 4 on a fair six-sided die, which is a non-zero value:

$$P(X = 4) = \frac{1}{6}$$

On the other hand,  $Y$  is a continuous random variable, meaning it can take on any value within a given range (in this case, between 1 and 6). For continuous random variables, the probability of taking on any specific value is zero. Therefore,

$$P(Y = 4) = 0$$

**b).** To calculate  $P(X > 4)$  for the discrete random variable  $X$ :

$$P(X > 4) = P(X = 5) + P(X = 6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

To calculate  $P(Y > 4)$  for the continuous random variable  $Y$ : The PDF of a uniform distribution  $U(a, b)$  is given by:

$$f_Y(y) = \begin{cases} \frac{1}{b-a} & a \leq y \leq b \\ 0 & \text{otherwise} \end{cases}$$

For our case,  $a = 1$  and  $b = 6$ , so the PDF is:

$$f_Y(y) = \begin{cases} \frac{1}{5} & 1 \leq y \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

To find  $P(Y > 4)$ , we can integrate the PDF from 4 to 6:

$$P(Y > 4) = \int_4^6 f_Y(y) dy = \int_4^6 \frac{1}{5} dy = \frac{1}{5} \cdot (6 - 4) = \frac{2}{5}$$

Thus,  $P(X > 4) = \frac{1}{3}$  and  $P(Y > 4) = \frac{2}{5}$ .

$$P(X > 4) = \frac{1}{3}, P(Y > 4) = \frac{2}{5}$$

**Problem 18.** Suppose  $X \sim U(0, 1)$ . Let a new random variable be defined as  $Y = -\ln(X)$ . Find the PDF of  $Y$ ,  $f_Y(y)$ . What famous distribution is this?

**Solution.** To find the PDF of  $Y$ , we can use the method of transformation. The CDF of  $Y$  is given by:

$$F_Y(y) = P(Y \leq y) = P(-\ln(X) \leq y) = P(X \geq e^{-y}) = 1 - P(X < e^{-y}) = 1 - e^{-y}$$

for  $y \geq 0$ .

Differentiating the CDF gives us the PDF:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = e^{-y} \text{ for } y \geq 0$$

This is the PDF of the exponential distribution with parameter  $\lambda = 1$ . Thus,  $Y$  follows an exponential distribution:

$$Y \sim \text{Exponential}(1)$$

**Problem 19.** A factory produces widgets, and 5% are defective. You take a random sample of 500 widgets. Use the normal approximation to the binomial distribution to estimate the probability that you will find 30 or more defective widgets in your sample. (Remember to use the continuity correction).

**Solution.** Let  $X$  be the number of defective widgets in the sample. Then  $X \sim \text{Binomial}(n = 500, p = 0.05)$ . The mean and variance of  $X$  are:

$$E[X] = np = 500 \times 0.05 = 25$$

$$\text{Var}(X) = np(1 - p) = 500 \times 0.05 \times 0.95 = 23.75$$

Using the normal approximation, we can approximate  $X$  with a normal distribution:

$$X \sim N(25, 23.75)$$

To find the probability of finding 30 or more defective widgets, we use the continuity correction:

$$P(X \geq 30) \approx P(X > 29.5)$$

We convert this to a z-score:

$$z = \frac{29.5 - 25}{\sqrt{23.75}} \approx \frac{4.5}{4.873} \approx 0.92$$

Using the z-table, we find:

$$P(Z < 0.92) \approx 0.81859$$

Thus, the probability of finding 30 or more defective widgets is:

$$P(X \geq 30) = 1 - P(Z < 0.92) \approx 1 - 0.81859 = \boxed{0.18141}$$

**Problem 20.** The time  $T$  (in minutes) to solve a complex puzzle follows an exponential distribution with a parameter  $\lambda = 1/15$ . If a player is paid based on their time according to the function  $P(T) = 100 - 2T$  dollars, what is the expected payout?

**Solution.** The expected payout can be calculated using the formula:

$$E[P(T)] = \int_0^{\infty} P(t) f_T(t) dt$$

where  $f_T(t) = \lambda e^{-\lambda t}$  is the PDF of the exponential distribution.

Given  $\lambda = \frac{1}{15}$ , the PDF is:

$$f_T(t) = \frac{1}{15} e^{-\frac{t}{15}} \text{ for } t \geq 0$$

Substituting  $P(t)$  and  $f_T(t)$  into the expected payout formula:

$$\begin{aligned} E[P(T)] &= \int_0^{\infty} (100 - 2t) \cdot \frac{1}{15} e^{-\frac{t}{15}} dt \\ &= \frac{1}{15} \int_0^{\infty} (100 - 2t) e^{-\frac{t}{15}} dt \\ &= \frac{1}{15} \left( 100 \int_0^{\infty} e^{-\frac{t}{15}} dt - 2 \int_0^{\infty} t e^{-\frac{t}{15}} dt \right) \\ &= \frac{1}{15} \left( 100 \left[ -15e^{-\frac{t}{15}} \right]_0^{\infty} - 2 \left[ -15te^{-\frac{t}{15}} - 225e^{-\frac{t}{15}} \right]_0^{\infty} \right) \\ &= \frac{1}{15} (100 [0 - (-15)] - 2 [(0 - 0) - (0 - 225)]) \\ &= \frac{1}{15} (100 \cdot 15 - 2 \cdot 225) \\ &= \frac{1}{15} (1500 - 450) \\ &= \frac{1050}{15} \\ E[P(T)] &= 70 \end{aligned}$$

Thus, the expected payout is 70 dollars.